

AD-A097 466

IOWA UNIV IOWA CITY DIV OF EDUCATIONAL PSYCHOLOGY F/G 5/10  
EXPERIMENTAL STUDIES OF CADA-BASED UTILITY ASSESSMENT PROCEDURE--ETC  
FEB 81 M R NOVICK, N J TURNER, L R NOVICK N00014-77-C-0428

UNCLASSIFIED

TR-81-2

NL

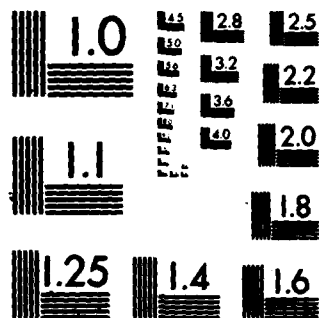
END

DATE

FILED

BY

DTIC



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

TECHNICAL REPORT 81-2

FEBRUARY, 1981

12  
LEVEL II

EXPERIMENTAL STUDIES OF CADA-BASED  
UTILITY ASSESSMENT PROCEDURES

MELVIN R. NOVICK, NANCY J. TURNER, AND LAURA R. NOVICK  
THE UNIVERSITY OF IOWA

AD A 097466

REPORT PREPARED UNDER OFFICE OF NAVAL RESEARCH

CONTRACT #N00014-77-C-0428

MELVIN R. NOVICK, PRINCIPAL INVESTIGATOR

THE UNIVERSITY OF IOWA

IOWA CITY, IOWA

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED  
REPRODUCTION IN WHOLE OR IN PART IS PERMITTED FOR  
ANY PURPOSE OF THE UNITED STATES GOVERNMENT.

FILE COPY

DTIC  
ELECTE  
S APR 8 1981 D  
D

81 4

7 001

(14) TR-81-2

Unclassified

(12) 75

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report 81-2 ✓	2. GOVT ACCESSION NO. AD A097466	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) "Experimental Studies of CADA-Based Utility Assessment Procedures"		5. TYPE OF REPORT & PERIOD COVERED Research Report. January - December 1980
6. AUTHOR(s) Melvin R. Novick; Nancy J. Turner and Laura R. Novick		7. PERFORMING ORG. REPORT NUMBER Technical Report 81-2
8. PERFORMING ORGANIZATION NAME AND ADDRESS University of Iowa Division of Educational Psychology Iowa City, Iowa 52242		9. CONTRACT OR GRANT NUMBER(s) N00014-77-C-0428 ✓
10. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research (Code 458) Arlington, VA 22217		11. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 6115 3N; RR 042-04; RR042-04-01; NR 150-404
12. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (16) RR04204 (17) RR0420401		13. REPORT DATE February 1, 1981
14. DISTRIBUTION STATEMENT (of this Report) Approved for public release, distribution unlimited		15. NUMBER OF PAGES 63
15. SECURITY CLASS. (of this report) Unclassified		16. DECLASSIFICATION/DOWNGRADING SCHEDULE
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Utility Assessment Coherence, fixed-state assessment, regional coherence, local coherence, CADA, certainty effect, anchoring effect		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Several experiments were performed to investigate the degree of incoherence associated with, and the extent of biases caused by, the <u>certainty</u> and <u>anchoring effects</u> for the following utility assessment procedures: standard gamble fixed state (without and with overfitting and least squares incoherence resolution), regional coherence, and local coherence. Several elicitation formats were also compared. The purpose of these experiments was to assist in the development of a reliable (coherent), bias-free, CADA-based method of utility assessment for general use in educational evaluation.		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 68 IS OBSOLETE  
S/N 0102-LF-014-6601

Unclassified

411417

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

January 30, 1981

Accession For	
DTIC GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

## EXPERIMENTAL STUDIES OF CADA-BASED

### UTILITY ASSESSMENT PROCEDURES<sup>1</sup>

Melvin R. Novick

Nancy J. Turner

Laura R. Novick

The University of Iowa

Several experiments were performed to investigate the degree of incoherence associated with, and the extent of biases caused by, the certainty and anchoring effects for the following utility assessment procedures: standard gamble fixed state (without and with overfitting and least squares incoherence resolution), regional coherence, and local coherence. Several elicitation formats were also compared. The purpose of these experiments was to assist in the development of a reliable (coherent), bias-free, CADA-based method of utility assessment for general use in educational evaluation.

Utility assessment can be treated as a topic in the classical psychometric specialty of scaling (e.g., Torgersen, 1958) or more formally as part of a theory of measurement (Krantz, Luce, Suppes, and Tversky, 1971). Novick and Lindley (1978, 1979) have dealt with the use of utility functions for applications in education and have advocated the use of the standard gamble (von Neumann and Morgenstern, 1953) elicitation procedure with the addition of coherence checking using overspecification and a least squares fit. In this procedure utilities are inferred from probability judgments offered by assessors.

---

<sup>1</sup>This research was supported under Contract #N00014-77-C-0428 between the Office of Naval Research and The University of Iowa. Opinions expressed herein are those of the authors and do not necessarily reflect those of the contracting agencies. The computer program development of David Libby and the consultative assistance of David Chuang are gratefully acknowledged.

Earlier approaches to utility assessment (Mosteller and Nogee, 1951; Schlaifer, 1959, 1971; Raiffa and Schlaifer, 1961; Keeney and Raiffa, 1976; and so on) have been based on the use of fixed probability assessment procedures in which utilities are elicited directly. It has been suggested (Mosteller and Nogee, 1951) that such procedures are easier to use because subjects are generally more familiar with the quantity for which the utility function is desired than they are with probabilities. Since use of an easier procedure does not necessarily insure that subject's responses will be coherent, an opportunity for incoherence resolution is needed. However, the nature of the fixed probability gambles makes it very difficult to check for coherence of responses.

Although it was originally thought that utility theory would prove useful as a descriptive model (Swalm, 1966, etc.), much criticism has recently been levied against the use of utility theory in that capacity (Coombs, 1975; Kahneman and Tversky, 1979). As principal critics, Kahneman and Tversky have proposed an alternative descriptive model. The main basis for their criticism is that the phenomenon described by Tversky (1977) as the certainty effect results in preferences that violate the substitution axiom or expected utility hypothesis of utility theory. This axiom (hypothesis) states that preference order is invariant over probability mixtures and is formally equivalent to the assumption that there is no positive or negative utility for the act of gambling itself. Specifically, the certainty effect is the phenomenon that the utility of an outcome seems greater when it is certain than when it is uncertain. This effect can be observed when subjects are presented with a choice between a for-sure and a chance option, the choice presented in the standard gamble, regional coherence, and local coherence assessment procedures discussed below.

Although utility theory is considered here as a normative model rather than as a descriptive model, it is still important to consider the certainty effect. Tversky (1977) has shown that even when subjects were told that their preferences violated utility theory, they were not inclined to change them (see also Kahneman and Tversky, 1972). This brings into question the reliability (coherence) and bias-free character of utility assessment procedures and the value of those procedures in helping decision makers be more coherent. However, two additional comments must be made. Whereas the gambles studied by Kahneman and Tversky involved gains and losses from some undefined reference point, those studied by Novick and Lindley considered passage to and from well-defined states. Furthermore, the latter authors also included incoherence resolution as a part of their utility assessment procedure.

In another paper, Tversky and Kahneman (1974) described several heuristics used by persons in assessing probabilities and the biases to which they could lead. Of particular interest is the anchoring and adjustment heuristic, whereby the most readily available piece of information often forms an initial basis for formulating responses, from which subsequent responses are then adjusted. Since adjustments from this basis are often insufficient, a central bias results. According to Slovic (1972), the anchoring and adjustment heuristic is a natural strategy for easing the strain of integrating information. The anchor serves as a register in which one stores first impressions or the results of earlier calculations. Slovic

advances two hypotheses to explain why adjustments from the anchor are usually insufficient. First, people may stop adjusting too soon because they tire of the mental effort involved in adjusting. Alternatively, the anchor may take on a special salience, causing people to feel that there is less risk in making estimates close to it than in making estimates far from it.

In reviewing the role of man-machine systems in decision analysis, Slovic, Fischhoff, and Lichtenstein (1977) suggested that human factors such as the ways in which variations in instructions or informational displays affect people's performance should be studied in more detail. Some progress has been made in this area with the finding that questions of complexity and representativeness of material seem to have substantial effect on assessors' responses (Fischhoff, Slovic, and Lichtenstein, 1977; Vlek, 1973). The study of such factors might lead to an assessment procedure that minimizes the judgmental biases and heuristics described earlier. This position was strengthened by the discussion of Fischhoff, Slovic, and Lichtenstein (1979).

The purpose of the experiments reported here was to assess the reliability (coherence) and degree of bias of several proposed utility assessment procedures. In particular, the influence of the certainty and anchoring effects identified by Tversky and Kahneman



(1977) was of interest. New response formats and recently proposed local and regional coherence (Novick, Chuang, and DeKeyser, 1981) procedures were considered. The ultimate goal of the experiments was to assist in the development of a reliable, bias-free, CADA-based utility assessment procedure for general use in educational evaluation.

Extensive previous work in this area has raised more questions concerning bias and coherence than it has provided answers. An apparently pessimistic mood prevails, not inappropriately, given the importance of the questions that have been raised. Nevertheless, the very extensiveness of this research must itself imply a high assessment for the product of the probability for resolving these difficulties and the value of this outcome. The position taken here is that bias and incoherence may be reduced if elicitations are carefully fashioned in a Computer-Assisted Data Analysis (CADA) environment (Novick, Hamer, Chen, Woodworth, Libby, Isaacs, Lewis, Molenaar and Chuang, 1980) and assessors are aided in resolving incoherence.

The outcomes for which utilities were sought were college grade point averages (GPAs) at graduation and average graduate record examination (GRE) scores of first-year graduate students. The following fixed state utility assessment procedures were compared:

1. Standard fixed state (SFS)
2. Regional coherence (RC)
3. Local coherence (LC)

The order of the steps taken by procedure 1 to elicit utilities (see description below) made possible an evaluation with and without incoherence resolution. The following formats were studied for obtaining indifference probabilities:

1. Direct probability elicitation
2. Ends-in method
3. Quartering method
4. Modified ends-in method with graphic representation

In addition, three sets of gambles were studied with the SFS procedure:

1. Adjacent gambles
2. Distant symmetric gambles
3. Utility gambles

Experimentation was sequential with conclusions drawn and insights gained from pilot experiments used to select methods and formats to be tested in the final experiment.

The three types of gambles refer to how the GPA's (alternatively, GRE scores) were chosen to construct those gambles. Adjacent gambles consisted of three successive GPAs at 0.5 intervals. That is, the three GPAs used in each gamble covered an interval equivalent to one point on the GPA scale--for example, (2.0, 2.5, 3.0). Distant symmetric gambles used successive GPAs at 1.0, 1.5, or 2.0 intervals. They were symmetric in the sense that the distances between the lowest and middle GPAs and between the middle and highest GPAs were equal--for example, (2.0, 3.0, 4.0) and (0.0, 1.5, 3.0). Utility gambles derive their name from the fact that the indifference probabilities elicited for each gamble are identical to the utilities for the for-sure GPAs of those gambles. This is achieved by setting the lower and upper GPAs in the chance option of each gamble at 0.0 and 4.0, respectively. The following are examples of utility gambles: (0.0, 1.0, 4.0) and (0.0, 3.5, 4.0). Experiment 1 was conducted in part to

determine which gambles in each set produced particularly large errors and thus should not be used subsequently.

In the SFS procedure subjects were given situations consisting of a for-sure and a chance option and were asked for the chance option probabilities that would make them indifferent with respect to the two options in each situation (i.e., their indifference probabilities). The indifference probabilities for the fixed state gambles (situations) could be elicited using any of four methods, corresponding to four formats for presenting the gambles, and the determination of the format which elicited the most coherent set of utilities was of primary concern. Table 1 below describes each format. The gamble used in the table can be abbreviated as (2.0, 2.5, 3.0) and this notation will henceforth be used to describe fixed state gambles.

Format One

FOR SURE OPTION			CHANCE OPTION		
I	GPA	I	I	GPA	I
I		I	I	3.5	PZ
I	2.5	I	I		I
I		I	I	1.5	100%-P

What value of p makes you indifferent between the for-sure and the chance options? \_\_\_\_\_

Format Two

FOR SURE OPTION			CHANCE OPTION			OPTION	
I	GPA	I	I	GPA	I	0	INDIFFERENT
I		I	I	3.5	P	1	FOR SURE
I	FOR SURE 2.5	I	I	CHANCE		2	CHANCE
I		I	I	1.5	1-P	3	RESTART

Which option would you prefer if p = .XX? \_\_\_\_\_

Format Three

FOR SURE OPTION			CHANCE OPTION			OPTION	
I	GPA	I	I	GPA	CHANCE	I	OPTION
I		I	I	3.5	75%	I	1 FOR SURE
I	2.5 FOR SURE	I	I			I	2 CHANCE
I		I	I	1.5	25%	I	3 RESTART

Which option do you prefer? \_\_\_\_\_

FOR SURE OPTION			CHANCE OPTION		
I	GPA	I	I	GPA	CHANCE
I		I	I	3.5	P
I	2.5 FOR SURE	I	I		
I		I	I	1.5	Q

Choose the set of P and Q values that will make you indifferent between the for-sure and the chance options.

P 75% 70% 65% 60% 55% 50%  
Q 25% 30% 35% 40% 45% 50%

The P value that makes you indifferent is ? \_\_\_\_\_

Table 1. Final versions of the formats for eliciting indifference probabilities

Format Four  
A

FOR SURE OPTION			CHANCE OPTION		
I	GPA	I	IGPA	I	I
I		I	I	I	I
I	2.5	I	I 1.5	I	I
I		I	I	I	I
100%			20%	80%	

- OPTION  
0 INDIFFERENT  
1 FOR SURE  
2 CHANCE  
3 RESTART

Which option do you prefer? \_\_\_\_\_

B

FOR SURE OPTION			CHANCE OPTION		
I	GPA	I	IGPA	GPA	I
I		I	I		I
I	2.5	I	I 1.5	3.5	I
I		I	I		I
100%			100%-P	P	

Your indifference probability, P, has been determined to be between 40% and 60%.

Value of P? \_\_\_\_\_

Table 1 Final versions of the formats for eliciting indifference probabilities  
con't.

Format One asked subjects to directly specify the value of  $p$  that will make them indifferent with respect to the for-sure and chance options. Hence it is known as the direct probability elicitation format. Originally subjects were asked to state probability values but in the final version of the format they stated percents instead. We found that subjects had difficulty thinking in terms of mathematical probabilities and were much more comfortable dealing with percents. Appendix A shows the evolution of this and other formats from our original conceptions of them to their final forms.

Format Two differs from Format One in that subjects were asked to state preferences between the two options given certain values for  $p$ . The computer asked for option preference or indifference regarding probabilities in the following order: .1, .9, .3, .7, .4, .6, .5 (alternatively, .9, .1, .7, .3, . . .). If the indifference probability was found to be between .6 and .7, say, then the questioning procedure continued with values of .62, .68, .64, . . . (.68, .62, .66 . . .). In this manner format two zeroed in on the subject's indifference probability for each situation. Because of the order in which the  $p$  values were presented, Format Two is known as the ends-in method.

Format Three, the quartering method, divided the probability scale, .05 - .95, into approximate quarters and attempted to determine in which quarter the subject's indifference probability lay. Subjects were first asked whether the probability that would make them indifferent between the two options was larger or smaller than .5. Depending on their answer to that first option, they were then asked whether their indifference probability was greater or smaller than either .25 or .75. The responses to these two questions placed the subject's indifference probability in one of the following four regions: .05 - .25, .25 - .50, .50 - .75, and

.75 - .95. Once the appropriate region was determined, subjects were presented with a list of probabilities at .5 intervals and asked to choose the set of probabilities (each set included  $p$  and  $1 - p$ ) that was closest to making them indifferent between the two options. Values for  $p$  below .5 were presented in ascending order (e.g., .05, .10, .15, .20, .25) and probabilities above .5 were presented in descending order (e.g., .75, .70, .65, .60, .55, .50). This was done in an attempt to avoid an anchor at .5.

The modified ends-in method with graphic representation, Format Four, was designed in an attempt to render the situations easier for unsophisticated subjects to understand, thus making it easier for them to respond. The field in which the chance option was displayed was divided into nine equal parts, thus allowing any of the probabilities from .1 to .9 at .1 intervals to be graphically portrayed. As in the original ends-in format, subjects were asked to state their preference (or indifference) between the for-sure and chance options as the values of  $p$  changed. Although originally  $p$  alternated as follows- .1, .9, .2, .8, .3, .7, .4, .6, .5 (alternatively, .9, .1, .8, . . .)- this procedure was found to be tedious and time-consuming. In the final version for Experiment 3 the number of  $p$  values subjects considered was drastically reduced (.2, .8, .4, .6 or .8, .2, .6, .4). If the indifference probability was found to be between .4 and .6 (.6 and .7 in the original version), say, the subjects were asked to directly specify probabilities of getting the higher and lower GPA's in the chance option that would make them indifferent between the two options. Note that whereas the above explanations were all given in terms of probabilities, Formats Three and Four and the final version of Format One were actually presented to the subjects in terms of percents (i.e., the probability values for the chance options were actually percents adding up to 100%).

The most coherent gambles (adjacent and distant symmetric gambles) and formats (Formats One and Four) were taken together to form the three fixed state utility assessment procedures for Experiment 3. Fifteen indifference probabilities were elicited for each subject. In the SFS procedure a nonlinear least squares (LSQ) fit of the data points was then made and subjects were subsequently allowed to re-assess their probabilities using their specified set as the working set. Subjects were permitted to re-specify their indifference probabilities a maximum of two times. It has been our experience that after this time few significant changes are likely to be made. If subjects did not wish to change their specified probabilities after the LSQ fit but were also not satisfied with the fitted probabilities, they were allowed to go directly to the questionnaire anyway. No subject was encouraged to change his/her indifference probabilities, only to think about the situations carefully. In this manner as much coherence as possible among the indifference probabilities, and thus the utilities, was obtained. Novick and Lindley (1978) have stressed the importance of obtaining coherence.



In the RC procedure indifference probabilities were elicited for two gambles using either Format One or Format Four. The subjects were then presented with a table showing the initial situations with their stated indifference probabilities and two additional situations with computer-deduced indifference probabilities. They were told that their initial responses implied certain specific indifference probabilities for the two new situations (situations three and four below). The following table illustrates the latter part of this procedure.

	Situations			
	-- 1 -----	2 -----	3 -----	4 ----
p chance	1.50	2.00	2.00	2.00
for sure	1.00	1.50	1.00	1.50
100%-p chance	0.50	1.00	0.50	0.50
	p=53%	p=58%	p=40%	p=75%

Table 2. RC Procedure

After studying the table the subjects were given an opportunity to change the indifference probabilities for two situations at a time until they expressed indifference with respect to the two options (chance and for-sure) in each of the four situations. In this way subjects were helped to make their probability judgments coherent for a region of grade point averages. When the subjects were satisfied with all four indifference probabilities, the procedure began again with specification of indifference probabilities for two new situations. A LSQ fit of the data points was done at the end of the RC procedure but was not shown to the subject. This allowed us to see how coherent the entire set of indifference probabilities was. Although the RC procedure determines that each group of four indifference probabilities is internally consistent, it does not guarantee that the probabilities will be consistent across groups.

The LC procedure presented subjects with two types of hypothetical choice situations: (1) a for-sure and a chance option (the standard gamble used in the SFS and RC procedures) and (2) two chance options. Again, either Format One or Format Four was used to elicit an indifference probability for the first situation, after which the subject was told that that response implied that he/she should be indifferent with respect to the two options in situation 2 for the designated probabilities. (The Option 1 probabilities for Situation 2 were always 50%/50% for the two highest GPAs as shown. Only the Option 2 probabilities for the highest and lowest GPAs changed.)

GAMBLE SITUATION 1				GAMBLE SITUATION 2			
I	OPTION 1	OPTION 2	I	I	OPTION 1	OPTION 2	I
I	(FOR SURE)	(CHANCE)	I	I			I
I		1.00 50%	I	I	50%	1.00 75%	I
I	0.50		I	I	50%	0.50	I
I		0.00 50%	I	I	—	0.00 25%	I
I			I	I			I

Table 3. LC Procedure

If the subject was not indifferent between the options in both situations, he/she was allowed to modify the Situation 1 indifference probability. After that probability was changed the subject was again presented with a table similar to Table 3 above, with the Situation 2 Option 2 probabilities now also different. This continued until the subject was indifferent with respect to the options in each of the situations. At that point the entire procedure was repeated with a new set of GPAs (i.e., a new gamble). As with the RC procedure, a LSQ fit of the indifference probabilities was computed after the subject had left.

An important feature of these experiments was that the primary interaction was between the assessor and a preprogrammed computer presentation of the experimental protocol. This presumably led to a high degree of uniformity in presentation. However, research assistants did read text and answer questions so it was possible that some inter-assistant variability might be present.

### Pilot Experiment 1 - Formats and gambles for the SFS procedure

Experiment 1 was designed to determine which of the first two formats, direct probability elicitation and ends-in, for presenting gambles in the three fixed state assessment procedures elicited the most coherent, bias-free set of indifference probabilities. The experiment was done in the context of the SFS procedure and the results were then extrapolated to the RC and LC procedures. Experiment 1 also determined which gambles resulted in the largest errors, in terms of log-odds differences, and thus should be omitted from the utility assessment procedures.

#### Method

Subjects. The subjects for the first part of experiment 1 were 43 college undergraduates of both sexes who responded to newspaper and other advertisements. The students were randomly assigned to four groups so that each group contained approximately the same number of subjects of each sex. Three subjects were excluded from the analyses because they failed to get through the entire program and thus contributed no data to the file. All subjects were paid \$4.00 for their participation in this experiment.

The subjects for the second phase of this experiment were PhD level professionals in the fields of Education, Psychology, Measurement, and Computer Science, and a few advanced graduate students, attending the Advanced Seminar in Bayesian Statistics supported by the CADA Research Group at The University of Iowa. These subjects served as volunteers.

Materials and Design. For all experiments Component 31 of the Computer-Assisted Data Analysis (CADA) Monitor (Novick, et al., 1980) was modified slightly and then used in a conversational mode to present the formats, gambles, and incoherence resolution methods associated with each assessment procedure. The Monitor was accessed via a cathode ray tube (CRT) terminal linked to a PDP 11V03 computer. Individual sessions were conducted by the research assistants who operated the computer system, recording verbal responses given by each participant. A complete recording of the protocol of each session was stored on flexible disk and made available in hard copy through a Decwriter terminal. Data analyses were performed on the 11V03 using data stored on the flexible disks.

The following two sets of gambles were used alternately with formats one and two: [1]--(0.0, 0.5, 1.0), (0.5, 1.0, 1.5), (0.0, 1.5, 3.0), (0.0, 1.5, 4.0), (1.5, 2.0, 2.5), (1.0, 2.0, 3.0), (0.0, 2.5, 4.0), (2.0, 3.0, 4.0), (0.0, 3.5, 4.0) and [2]--(0.0, 0.5, 4.0), (0.0, 1.0, 2.0), (0.0, 1.0, 4.0), (1.0, 1.5, 2.0), (0.5, 2.0, 3.5), (2.0, 2.5, 3.0), (1.5, 2.5, 3.5), (0.0, 3.0, 4.0), (3.0, 3.5, 4.0). For the professional-level subjects, average GRE scores rather than GPAs were used. To construct the gamble sets, 400 was substituted for 0.0, 450 for 0.5, etc. up to 800 for 4.0. The nine gambles comprising each set were chosen so as to make the two sets of gambles as similar as possible. Even so, both formats were used with each set of gambles. The subjects in group one responded to the first set of mixed gambles using Format One and then the second set using Format Two. Group two subjects responded first to set one using Format Two and second to set two using Format One. Groups three and four responded to the set two gambles first, using Formats One and Two respectively. The design was thus a crossed formats by gambles design with a further administration of the alternate format-gamble pair to each group. A resolution program was used to determine the best format, in the sense of minimum mean squared deviations from a least squares fit over all gambles made by the participant. (See data analysis section.) This design thus permitted examination of the format effect with a check and adjustment for gamble-set and order effects.

Procedure. The experiment was conducted in a private office with subjects being tested individually in one-hour sessions. Each subject participated in one session. All instructions were standardized and printed on a CRT terminal. The research assistant read the instructions to the undergraduates to insure that they clearly understood what was expected of them. The professional-level subjects, however, were permitted to read the instructions on their own, with the research assistant there to answer any questions they may have had. Subjects were given very careful instructions in the meaning of indifference probabilities and in the assessment of them. The concept of utility was never mentioned, nor were a subject's utilities ever displayed, because of the difficulty of making utility theory understandable to people with little or no statistical background. Some of the professional-level subjects were familiar with Bayesian statistics and utility theory and this was duly recorded. Except for the exact sets of instructions printed, all of the above was standard procedure in both of the pilot experiments.

In addition to the general instructions, specialized instructions were given for each elicitation format. Prior to indicating their first indifference probability, subjects were given a chance to review the instructions to insure that they clearly understood what they were to do. After the introductory paragraph and before the instructions concerning the nature of the choice situations, the following four questions were asked of undergraduates to get them to begin considering how their final grade point average would affect their future opportunities:

1. WHAT YEAR IN SCHOOL ARE YOU?
2. WHAT IS YOUR MAJOR?
3. UPON COMPLETION OF YOUR BACHELOR'S DEGREE, WHICH OF THE FOLLOWING POSSIBILITIES DO YOU CONSIDER MOST LIKELY TO OCCUR?

GRADUATE STUDY TO MASTER'S DEGREE LEVEL  
GRADUATE STUDY TO PHD LEVEL OR PROFESSIONAL SCHOOL  
EMPLOYMENT  
OTHER  
ANSWER?

4. TO WHAT EXTENT DO YOU FEEL YOUR CUMULATIVE GPA WILL AFFECT WHAT YOU WANT TO DO IN THE FUTURE?

GREATLY  
SOMEWHAT  
VERY LITTLE  
ANSWER?

In order to get the professional-level subjects to begin thinking about the importance of GRE scores in selecting first-year graduate students, the following four questions were asked of them:

1. WHAT IS THE ACADEMIC DISCIPLINE FOR WHICH YOU MIGHT SERVE AS CHAIR OF GRADUATE ADMISSIONS?
2. WHAT IS THE APPROXIMATE MEAN AVERAGE GRE SCORE OF FIRST-YEAR GRADUATE STUDENTS IN THE DEPARTMENT IN WHICH YOU MIGHT SERVE? RECALL THAT EACH SUBTEST OF THE GRE IS ON THE SCALE 200-900 WITH A NATIONAL MEAN OF ABOUT 500 AND STANDARD DEVIATION OF ABOUT 100.  
ANSWER?

3. DO YOU FEEL THAT THIS SCORE IS A GOOD PREDICTOR OF FUTURE PERFORMANCE IN YOUR PROGRAM?

1. YES, VERY GOOD PREDICTOR
2. YES, MODERATELY GOOD PREDICTOR
3. AVERAGE PREDICTOR
4. NO, FAIR PREDICTOR
5. NO, POOR PREDICTOR

ANSWER?

4. DO YOU FEEL THAT THE TEST SCORE IS AN IMPORTANT PIECE OF INFORMATION IN SELECTING STUDENTS?

1. YES, VERY IMPORTANT
2. YES, IMPORTANT
3. NO STRONG FEELINGS EITHER WAY
4. NO, NOT VERY IMPORTANT
5. NO, NOT IMPORTANT AT ALL

ANSWER?

A complete set of the instructions given to these subjects may be found in Appendix B. Except for a few slight modifications to make the instructions easier to understand, the same instructions were used for the professional-level subjects as for the undergraduates. Thus to derive the instructions for the undergraduates, one can simply substitute "final grade point average at graduation" for "admittance of a student with a particular average GRE score." The paragraphs introducing the scenario were different for both sets of subjects and thus both introductions are reprinted in Appendix B.

After providing indifference probabilities for both sets of gambles, subjects were shown a least squares fit of the data points. They were told that the computer had generated a set of indifference probabilities similar to theirs in order to help them specify a coherent set of  $p$  values. If subjects were not satisfied with the fitted probabilities, they were given an opportunity to change any or all of their specified probabilities. A new least squares fit based on the revised probabilities was then computed. Subjects were allowed to revise their indifference probabilities twice if they so desired and thus were able to see up to three different sets of fitted probabilities. Any subject who had not accepted the fitted probabilities by the third time was asked to do so then because it has been our experience that by then neither set of probabilities is likely to change substantially. When the subject accepted the fitted probabilities the formal part of the experiment was complete.

At the end of the experiment all subjects responded to a questionnaire calling for an evaluation of the experiment as a whole. If subjects inquired further about the purpose of the experiment they were told that the results would be used to help design procedures for making it easy for persons to decide on desirable gambles.

## Pilot Experiment 2 - Formats for standard fixed state gambles

Experiment 2 was an extension of experiment 1; its purpose was to determine which of two new formats (formats 3 and 4) for eliciting indifference probabilities provided the most coherent, bias-free set of probabilities. A revised set of gambles was used.

### Method

Subjects. The 10 subjects were graduate students enrolled in the Bayesian Statistics I class at The University of Iowa. They served as volunteers.

Design. The design was a crossed formats by gambles design with a further administration of the alternate format-gamble pair to each group as in experiment 1. Gamble set one consisted of the following adjacent gambles: (0.5, 1.0, 1.5), (1.0, 1.5, 2.0), (1.5, 2.0, 2.5), (2.0, 2.5, 3.0), and (2.5, 3.0, 3.5). Gamble set two included the following distant symmetric (specifically, two-apart) gambles: (0.0, 1.0, 2.0), (0.5, 1.5, 2.5), (1.0, 2.0, 3.0), (1.5, 2.5, 3.5), and (2.0, 3.0, 4.0). In addition, whichever set of gambles was responded to first, regardless of format, also included (0.0, 0.5, 1.0) and (3.0, 3.5, 4.0) so that the first seven situations were sufficient to determine a utility function. The second set of gambles to which a subject responded included (0.0, 2.0, 4.0) and (1.0, 2.5, 4.0). Thus each subject indicated indifference probabilities for a total of 14 situations. Computation of format errors did not consider the above four gambles. These 14 gambles were chosen on the basis of the results of Pilot Experiment 1. The utility gambles were excluded because subjects found them difficult and because they generally resulted in large log-odds differences between the assessed and fitted indifference probabilities. The four experimental groups were constructed as in the previous experiment.

Procedure. The procedure was the same as in experiment 1 except that each subject indicated only 14 indifference probabilities rather than 18.

### Experiment 3 - Comparison of SFS, RC, and LC procedures

The purpose of this experiment was twofold. First we were interested in whether the original format (Format One) or a new format (Format Four) resulted in the most reliable (coherent) and bias-free set of utilities (indifference probabilities). The major conclusion of interest, however, was which of the three utility assessment procedures provided the best set of utilities.

#### Method

Subjects. Although 50 subjects were run, 10 had to be excluded from the data analyses because a) they failed to get through the entire procedure or b) their indifference probabilities were too incoherent for a least squares fit to be computed. The remaining 40 were randomly assigned to three groups, with approximately the same number of subjects in each group. Subjects were paid \$4.00 upon completion of the experimental session. Students who had participated in either of the previous pilot experiments were not allowed to be in this experiment.

Design. In this experiment format was a within-subject variable while assessment procedure and format order were between-subject variables. Each group of subjects used only one of the three procedures. Half of the subjects in each group used Format One followed by Format Four, while the other subjects used Format Four first. There were three or four subjects of each sex in each format-order group for each procedure.

Format One, direct elicitation, was chosen because it is the format most often used with utility assessment procedures. Format Four, on the other hand, was chosen because Pilot Experiments 1 and 2 suggested that it may be better than Format One, in the sense of being easier for assessors to use, in inducing more coherent responses from them, and in



seemingly being less influenced by the anchoring effect. The gambles were chosen on the basis of the same information as led to the format choices and, in general, were the ones subjects found easiest to think about. The only utility gamble used is (0.0, 2.0, 4.0) because subjects generally found it very difficult to specify indifference probabilities for utility gambles. As a group, the utility gambles led to much more incoherent judgments than either the adjacent or distant symmetric gambles. There were a few other gambles (e.g. (0.0, 0.5, 1.0) and (3.0, 3.5, 4.0)) which were a little difficult for subjects but these were retained because they were needed to compute the utility function. Also, these gambles were not as bad as the utility gambles.

Two mixed sets of gambles (containing both adjacent and distant symmetric gambles) were constructed. The seven gambles in Set One were sufficient to compute a utility function and were used as the basis for determining the fitted probabilities for the SFS procedure. Set Two included all of the remaining adjacent and distant gambles. The gamble sets were:

SET ONE - (0.0, 0.5, 1.0), (0.0, 1.0, 2.0), (0.5, 1.5, 2.5), (1.5, 2.0, 2.5), (1.0, 2.5, 4.0), (2.5, 3.0, 3.5), and (3.0, 3.5, 4.0).

SET TWO - (0.5, 1.0, 1.5), (1.0, 1.5, 2.0), (0.0, 1.5, 3.0), (1.0, 2.0, 3.0), (0.0, 2.0, 4.0), (2.0, 2.5, 3.0), (1.5, 2.5, 3.5), and (2.0, 3.0, 4.0). These two sets were used as presented above for both the SFS and LC procedures. Because of the manner in which situations three and four were generated in the RC procedure, it was necessary to substitute (0.0, 1.0, 2.0) and (2.5, 3.0, 3.5) for (0.0, 1.5, 3.0) and (2.0, 2.5, 3.0) in Set Two for that procedure. Thus subjects responded to these two gambles twice in the RC procedure but only once in the other two procedures.

Subjects always responded to the Set One gambles before proceeding to the Set Two gambles. The formats, however, as explained above, were counterbalanced with half of the subjects using each procedure receiving Format One-Set One/Format Four-Set Two and the other half receiving Format Four-Set One/Format One-Set Two. This was done in order to enable us to look at format effects within and across procedures. The design was thus a crossed procedure by sex by format-order design. Gamble set order, however, remained fixed.

Procedure. The procedure for this experiment was in general the same as that for the two pilot experiments. There were, however, a few differences. First, the exact procedure followed by each subject was obviously dependent on whether he/she used the SFS, RC, or LC procedure. In the previous studies all subjects used the SFS procedure. Also, in this experiment we had subjects write down their utility function for grade point averages (from 0.0 to 4.0) on a scale of 0-100 after they had completed the probability elicitation procedure. This question was in addition to the general written questionnaire.

Data Analysis. Ideally a data analysis would compare assessed utilities with true utilities. Unfortunately there is no assured method of measuring true utilities. However, an approach commonly adopted in psychological testing seems relevant. Assuming that assessed utilities involving extensive measurements with averaging and/or adjustment are likely to be close to true utilities, the reliability or accuracy of individual component procedures can be assessed by comparing utilities from these procedures with those from the more extensive measurement. The testing analogy would be to evaluate each of several short measures of a construct by relating each measure to the average overall measures or perhaps to some factor analytically defined composite.

A variety of methods of comparing measures is possible. We have chosen to use an approach that is closely related to the least squares fitting procedure (Novick & Lindley, 1979). Consider, for example, the comparison of two methods (I and II), each involving  $m$  fixed state assessments for  $N$  points, with coherence adjustment. Denote the  $N$  utilities obtained for the two methods of  $U_{I1}, \dots, U_{IN}$  and  $U_{II2}, \dots, U_{IIN}$ . Denote the  $N$  assessed utilities after the least squares fit and resolution involving the  $2m$  gambles by  $U_1, \dots, U_N$ . The symbols  $U_I, U_{II}$ , and  $U$  will be used to denote vectors containing these ordered utilities. For each gamble used there will be a probability implied by each of the vectors  $U_I, U_{II}$ , and  $U$ . We denote these as value  $\pi_{I1}, \dots, \pi_{Im}, \pi_{II1}, \dots, \pi_{IIm}$ , and  $\pi_1, \dots, \pi_m$ , and the corresponding vectors by  $\pi_I, \pi_{II}$ , and  $\pi$ . Then taking the log odds ( $\gamma = \log [\pi/(1-\pi)]$ )

of the individual probabilities we can evaluate each method by considering discrepancies between transformed probabilities for the individual methods and the corresponding values for the final fitted utilities. In these analyses we shall take the values  $\gamma_1, \dots, \gamma_N$  for the resolved log-odds probabilities as "true" values and consider values  $\gamma_{I1}, \dots, \gamma_{IN}$  and  $\gamma_{III1}, \dots, \gamma_{IIIN}$  as measurements containing independent, normal, homoscedastic errors. The symbols  $\gamma, \gamma_I,$  and  $\gamma_{II}$  will denote the vectors of log-odds values. Data analysis will then involve the comparison of the average bias and average squared error for the methods in question. Symbolically the comparisons are  $(\gamma_{I1}-\gamma_1), (\gamma_{I2}-\gamma_2), \dots, (\gamma_{IN}-\gamma_N)$  and  $(\gamma_{III1}-\gamma_1), (\gamma_{III2}-\gamma_2), \dots, (\gamma_{IIIN}-\gamma_N)$ . In the several experiments this involves comparing various categories of method, format, type of gamble, and individual gambles.

The normality and homoscedasticity assumptions are reasonable on theoretical grounds by virtue of the log-odds transformation and are supported by data exploration. Slightly less comforting is the assumption that errors are independent (Novick & Lindley, 1979). However, measurement independence is a standard casual assumption that seems to be as acceptable here as in other applications. Failure of this assumption would be important primarily if it were differential between methods being compared.

## Results

Two dependent variables, Mean Absolute Error (MAE) and Greatest Absolute Difference in Utilities (GADU) were considered as criteria of subjects' performance. These measures and the relationship between them will be discussed in detail below. The tables presenting the results to be discussed are located in Appendix D.

### Analysis with MAE as the Dependent Variable

The primary criterion of a subject's performance was Mean Absolute Error, an index of overall consistency. Three types of MAE measures were utilized in the analysis. The overall measure for each subject was obtained by averaging the 15 absolute log-odds differences between his/her specified probabilities and the probabilities obtained from the least-squares fit (see Table 4). The distribution of MAE was determined to be approximately normal by CADA's Normal Probability Plot. The grand mean and the standard deviation for the measure were .4065 and .1800, respectively, with observations ranging from .0555 (very consistent) to .8540 (very inconsistent).

Because of the interest in performance under each of the two formats, two additional MAEs were computed for each subject: one for performance using Format 1 and one for performance using Format 4. The cell means and n-counts for the analysis of the effects of Procedure, Format, and Sex are presented in Table 5.

A final measure of MAE was considered appropriate for some aspects of the analysis. This variable may be described as the difference between the MAE obtained under Format 1 and the MAE obtained under Format 4 ( $\Delta \text{MAE}_{1-4}$ ). It was hypothesized that  $\text{MAE}_1$  would be somewhat greater than  $\text{MAE}_4$ , and so a small positive value was expected for each individual.

As can be seen from Table 6, the grand mean for  $\Delta MAE_{1-4}$  was slightly positive ( $M = .0046$ ), although there were many exceptions for the individual observations and also for several of the cell means.

A two-way ANOVA was applied to the new variable  $\Delta MAE_{1-4}$  in order to determine the extent to which it was affected by Procedure and Sex. Table 6 shows the marked Procedure effect upon  $\Delta MAE_{1-4}$  (39% of the total variance in  $\Delta MAE_{1-4}$  can be accounted for by Procedure). In particular, the average  $\Delta MAE_{1-4}$  for the RC procedure was relatively high (.0587), indicating a high degree of interaction between this procedure and the formats. The LC and SFS procedures evidenced lesser interactions with the formats (average  $\Delta MAE_{1-4}$  for SFS =  $-.0242$ , and average  $\Delta MAE_{1-4}$  for LC =  $-.0185$ ). These results are discussed in more detail below.

#### Main Effects for MAE

##### Procedure

While it was difficult to have prior convictions about the magnitude of the Procedure effects, the authors did have preconceptions about the efficacy of the three assessment procedures. In particular, it was expected that the Regional Coherence Procedure would be most successful in eliciting consistent responses, that the Standard Fixed State Procedure would be least effective, and that the Local Coherence Procedure would be intermediate. However, as seen in Table 4, these expectations were only partly realized. The MAE for RC, SFS, and LC were .3923, .4021, and .4295, respectively. The poor performance with the LC procedure was thought to be due to the lack of sophistication of the undergraduate subjects and might not be replicated with a sample of subjects who had had more statistical training. A procedure for making the LC procedure more useful might

consist of 1) revising the presentation to make it more comprehensible to neophytes, 2) rewriting instructions to make its use more attractive, or 3) using it with only more sophisticated subjects.

#### Format

The effect of Format on MAE was surprisingly small (see Table 5). We had expected that Format 4 would yield more consistent results. Format may be considered a within subject factor since each subject was run under both formats, and so 40 observations were obtained for each of the two formats. The average MAE under Format 4 was .4040, which could not be considered to be appreciably different from the MAE of .4103 obtained under Format 1. The Format factor might, however, be considered interesting due to its interaction with Procedure (to be discussed below).

#### Sex of Subject

The subjects' sex was found to have a strong effect on performance as measured by MAE. The greater consistency on the part of the female subjects was interesting for several reasons. First of all, it was expected that males would have a more extensive background in quantitative disciplines and might therefore excel in tasks requiring complex reasoning abilities. Secondly, there was a problem with students who had made appointments for the experiment failing to report, and this "no-show" tendency was markedly more prevalent in male subjects. Therefore, it was expected that the males who did keep their appointments would be

more highly motivated and more likely to perform well. While neither of these hypotheses was borne out by the MAE data, alternative explanations for the greater female consistency might be offered. In particular, the results could be seen as consistent with the general finding that women are more careful than men on many tasks. They might, therefore, be expected to perform better in an experiment which requires great deliberation and painstaking accuracy.

### Interaction Effects

#### Procedure by Format Interaction

As discussed above, Procedure had a strong effect on the variable  $MAE_{1-4}$ , accounting for 39% of its variance. A complex interaction between Format and Procedure was particularly apparent under the RC procedure (see Table 5).

Since one of the objectives of the experiment was to determine which combination of Procedure and Format would yield the most consistent results, particular attention should be paid to the average MAEs presented in Table 5. It should be noted that the RC procedure used in conjunction with Format 4 yielded the lowest average MAE (.3615). This finding confirmed expectations of both the relative efficacy of the RC procedure and the relative ease of Format 4. It might also be noted that this winning combination was particularly effective for the 6 female subjects who achieved very consistent results ( $\bar{x} = .2846$ ) through its use. On the other hand, the LC procedure coupled with



Format 4 produced the least desirable results ( $\bar{x} = .4387$ ), although these results did not differ greatly from those obtained under the LC procedure in conjunction with Format 1 ( $\bar{x} = .4202$ ).

#### Procedure by Sex Interaction

Table 4 demonstrates that males performed best under the LC procedure while female performance was worst under this procedure. The females apparently profited from the RC and SFS procedures (achieved means of .2978 and .3285, respectively) while the males did rather poorly under these two procedures (with mean performance of .4803 and .5046, respectively).

#### Format by Sex Interaction

Table 5 indicates little interaction between Format and Sex. Males performed only somewhat better under Format 4 than under Format 1 while there was essentially no difference in female performance as a function of format.

#### Analysis with GADU as the Dependent Variable

A second criterion of performance was the Greatest Absolute Difference in Utilities (GADU) which may be described as an index of congruence between the subject's stated utility function (written on a piece of paper before or after the experiment) and the utility function obtained by his specification of probabilities during the procedure. It was believed that this might be a better indicator of how well the subject was able to use the procedure to reproduce his GPA gambling behavior than would the consistency index (MAE) described above. This assumption was based on the observation of cases in which subjects appeared to quite consistent (perhaps merely by specifying .5 or some other value for most of the situations), but had fitted utility functions which deviated greatly from their

stated functions.

The hypothesis that consistency has little or no relationship to success in reproducing utility functions was confirmed by obtaining a Spearman Rank-Order Correlation coefficient for the ranks of each of 40 observations on overall MAE and GADU. The  $r_s$  of  $-.0858$  points to the fact that there is essentially no correlation between these two skills and also helps to explain the reversals in results when observations for these two variables were broken down by Procedure, Sex, and Format (see analysis for MAE above and GADU below). This lack of correlation may cast doubt on one or both of the indices as measures of successful completion of the gambling task or it may be a function of the rather small undergraduate sample tested. On the other hand, the essentially zero correlation may indicate that both indices are valid measures of performance but that they identify very different skills.

For the purposes of this analysis, the 40 subjects were divided into two groups depending on the format they used first in the utility assessment procedure. This was done for the following reason: The 15 gambles used in this experiment were divided into two sets. Set 1, containing 7 gambles, was always responded to first while Set 2, consisting of 8 gambles, was always responded to second. Since the for-sure options in the first gambles contained all of the GPAs between 0.5 and 3.5 (at 0.5 intervals), a utility function could be computed from the indifference probabilities for Set 1. However, this was not possible for the Set 2 gambles. Thus, for some of the subjects the fitted utility function was based on performance under Format 1

while for the remainder of the subjects the function was based on performance under Format 4. Consequently, there appear to be half as many subjects per format in Table 7 than in Table 5.

The 40 observations on GADU are presented in Table 7 along with the twelve cell means when GADU was broken down by Procedure, Format, and Sex. Table 8 provides summary data. The observations on GADU were essentially normally distributed with a mean of .3384 and a standard deviation of .18. They ranged from .1153 (very close approximation of stated utility function) to .8104 (very poor approximation).

#### Main Effects for GADU

##### Procedure

The Procedure effect is particularly interesting because of the ordering of the effectiveness of the procedures in eliciting a small GADU. As Table 7 indicates, the RC procedure produced the smallest GADU. The LC and SFS procedures were clearly less effective in obtaining utilities which were similar to the students' stated utility functions. It might be claimed that the checking devices offered by the two coherence procedures were helpful to subjects in the sense of making the probabilities upon which their utilities were based conform to the GPA utilities they had in mind. It should also be noted that the RC procedure led to better results on both the MAE and the GADU criteria which, as mentioned above, are hypothesized to measure two distinct skills.

##### Format

The Format effect was once again very small, as demonstrated by Tables 7 and 8. Since utilities were computed only from the first

seven gambles, only one measure of GADU was obtained for each subject, thereby permitting the three-factor analysis summarized in the two tables. The mean GADU for the subjects in the Format 4 group was only slightly lower than that for the subjects in the Format 1 group ( $\bar{x}_{F4} = .3448$  vs.  $\bar{x}_{F1} = .3597$ ).

#### Sex of Subject

The subject's sex once again had an important effect on the criterion variable. Perhaps, the most interesting observation that may be made here is that males (with a mean GADU of .3152) were clearly superior to females (with a mean GADU of .3836) in minimizing the difference between their stated and fitted utility functions (See Table 7). This contrasts with the superior female performance on obtaining consistent results as revealed by the MAE index.

#### Interaction Effects

##### Procedure by Format Interaction

The complexity of this interaction for dependent variable GADU was somewhat greater than for MAE (see Table 8A). For the RC Procedure, performance did not differ as a function of format. The LC procedure seemed better in conjunction with Format 1 than Format 4, while the reverse was true for the SFS procedure.

##### Procedure by Sex Interaction

A small interaction was noted between Procedure and Sex, as indicated in Table 8C. The results show that performance under the LC procedure evidenced the greatest variation as a function of sex. The GADU results differ from the MAE results in that GADU shows RC to be best for both males and females while MAE indicates that females

perform best under RC and males perform best under LC.

#### Sex by Format Interaction

This interaction was very slight. Females performed only somewhat better under Format 4 than under Format 1 while there was essentially no difference in male performance as a function of format. GADU and MAE were consistent in showing that performance under each format did not differ as a function of the sex of the subject.

#### Procedure by Format by Sex Interaction

The three-way interaction was large and striking as illustrated in Figure 1. The effect of format varied as a function of sex and utility assessment procedure.

#### Effect of Statistical Training on Performance

It was hypothesized that students who had had some statistical training would perform better than those who had had none. Indeed, the 14 students who had had one or more hours of statistics obtained an average MAE of .3411 while those with zero hours of statistics obtained an average MAE of .4012. When GADU was the criterion, the students with statistical training obtained an average score of .3061 while those with no training obtained an average score of .3583.

#### General Discussion

Several conclusions may be drawn from the three experiments conducted. First, adjacent and distant symmetric gambles appear to be more useful than utility gambles. Second, the Regional Coherence procedure was found to lead to superior results on both the MAE and GADU criteria. Finally, neither Format provided utilities more congruent

with initially stated utilities than the other. No conclusion could be drawn as to whether the fault lay in the initially stated or experimentally elicited utilities.

When taken together, however, the Regional Coherence procedure and the modified ends-in format (Format 4) seem promising, at least for people with moderate to no statistical training. Unfortunately, while the undergraduates participating in this experiment came from a wide range of academic disciplines, they may have lacked either the ability or motivation to fully appreciate and understand the complexity and importance of the gambling situations. A more statistically inclined sample of people in another environment might well get the best results with a different combination of procedure and format.

While even the trained statistician could no doubt benefit from the use of a utility assessment procedure with coherence-checking capabilities, the "naive" person would seem to gain the most from use of such a procedure in the interactive environment of CADA. And for these persons Experiment 3 suggests that Regional Coherence with a graphic ends-in method of eliciting indifference probabilities, if not the ultimate answer, is at least somewhat more profitable than its rivals. We feel confident that this combination or some derivative, when used to assess utilities in a real-life situation, will prove useful.

Format 1 - Direct Elicitation

1. As originally on CADA:

FOR SURE	GAMBLES		P THAT MAKES YOU INDIFFERENT
	P	I-P	
0.50	1.00	0.00	? .5
1.00	1.50	0.50	? —

2. As for Pilot Experiment 1A:

I	GPA	GPA	I
I		P 1.0	I
I			I
I FOR SURE	0.5		GAMBLE I
I			I
I	1-P	0.0	I

What value of P makes you indifferent between the for-sure and the chance options? \_\_\_\_\_

3. As for Pilot Experiment 1B:

Format One

I	GPA I	I	GPA	I
I	I	I	3.5	P I
I FOR SURE	2.5 I	I	CHANCE	I
I	I	I	1.5	1-P I

What value of p makes you indifferent between the for-sure and the chance options? \_\_\_\_\_

4. As for Experiment 3:

Format One

FOR SURE OPTION			CHANCE OPTION		
I	GPA	I	I	GPA	I
I		I	I	3.5	PX
I	2.5	I	I		I
I		I	I	1.5	100%-P

What value of p makes you indifferent between the for-sure and the chance options? \_\_\_\_\_

Format 2 - Ends-In

1. As originally on CADA:

I		P	1.00	I	OPTION:
I				I	0 INDIFFERENT
I	FOR SURE	0.50		I	1 FOR SURE
I			GAMBLE	I	2 CHANCE
I		1-P	0.00	I	3 RESTART

Which would you prefer if  $p = .XX$  ? \_\_\_\_\_

2. As for Pilot Experiment 1A:

Format Two

I		GPA	I	I		GPA	I	OPTION
I			I	I		3.5	P	1 FOR SURE
I	FOR SURE	2.5	I	I	CHANCE			2 CHANCE
I			I	I		1.5	1-P	3 RESTART

Which option would you prefer if  $p = .XX$ ? \_\_\_\_\_

Format 3 - Quartering

1. Original Conception:

I		GPA	I	I		GPA	I	OPTION
I			I	I		3.5	75%	1 FOR SURE
I	FOR SURE	2.5	I	I	CHANCE			2 CHANCE
I			I	I		1.5	25%	3 RESTART

Which option do you prefer? \_\_\_\_\_

I		GPA	I	I		GPA	I
I			I	I		3.5	P
I	FOR SURE	2.5	I	I	CHANCE		
I			I	I		1.5	Q

Choose the set of P and Q values that will make you indifferent between the for-sure and the chance options.

P 75% 70% 65% 60% 55% 50%

Q 25% 30% 35% 40% 45% 50%

The P value that makes you indifferent is? \_\_\_\_\_



2. As for Pilot Experiment 2:

Format Three

FOR SURE OPTION			A	CHANCE OPTION		
I	GPA	I		I	GPA	CHANCE
I		I		I	3.5	75%
I	2.5	FOR SURE	I	I		
I		I		I	1.5	25%

- OPTION  
1 FOR SURE  
2 CHANCE  
3 RESTART

Which option do you prefer? \_\_\_\_\_

B

FOR SURE OPTION				CHANCE OPTION		
I	GPA	I		I	GPA	CHANCE
I		I		I	3.5	P
I	2.5	FOR SURE	I	I		
I		I		I	1.5	Q

Choose the set of P and Q values that will make you indifferent between the for-sure and the chance options.

P 75% 70% 65% 60% 55% 50%  
Q 25% 30% 35% 40% 45% 50%

The P value that makes you indifferent is ? \_\_\_\_\_

Format 4 - Modified Ends-In

1. Original Conception:

Format Four

100%			A	20%			80%			OPTION
I	GPA	I		I	GPA	I	GPA	I		0 INDIFFERENT
I		I		I	I	I		I		1 FOR SURE
I	2.5	I		I	1.5	I	3.5	I		2 CHANCE
I		I		I	I	I		I		3 RESTART

FOR SURE

CHANCE

Which option do you prefer?

		B			
100%		100%-P		P	
I	GPA	I	IGPA	GPA	I
I		I	I		I
I	2.5	I	11.5	3.5	I
I		I	I		I
FOR SURE OPTION		CHANCE OPTION			

Your indifference probability, P, has been determined to be  
between 40% and 60%.

Value of P? \_\_\_\_\_

2. As for Pilot Experiment 2 and for Experiment 3:

Format Four  
A

FOR SURE OPTION			CHANCE OPTION		
I	GPA	I	IGPA I	GPA	I
I		I	I I		I
I	2.5	I	I1.5 I	3.5	I
I		I	I I		I
100%			20%	80%	

OPTION  
0 INDIFFERENT  
1 FOR SURE  
2 CHANCE  
3 RESTART

Which option do you prefer? \_\_\_\_\_

B

FOR SURE OPTION			CHANCE OPTION		
I	GPA	I	IGPA	GPA	I
I		I	I		I
I	2.5	I	I1.5	3.5	I
I		I	I		I
100%			100%-P	P	

Your indifference probability, P, has been determined to be between 40% and 60%.

Value of P? \_\_\_\_\_

**PILOT EXPERIMENT 1****A. Introduction:****experiment 1A**

THE PURPOSE OF THIS EXPERIMENT IS TO STUDY GAMBLING BEHAVIOR WHEN THE OBJECT OF THE GAMBLE IS AN IMPORTANT FACTOR RELATING TO A COLLEGE STUDENT'S FUTURE OPPORTUNITIES, NAMELY, CUMULATIVE GRADE POINT AVERAGE (GPA) AT GRADUATION.

YOU WILL BE PRESENTED WITH SEVERAL SITUATIONS INVOLVING CHOICES BETWEEN GAMBLING AND SURE THINGS. YOU ARE TO IMAGINE THAT THE OUTCOME OF THE OPTION YOU CHOOSE WILL RESULT IN THE FIXING OF A PARTICULAR GPA ON YOUR PERMANENT UNIVERSITY RECORD.

WHILE WE REALIZE THAT YOU ARE NOT ACCUSTOMED TO GAMBLING ON YOUR GRADE POINT AVERAGE, WE WOULD LIKE YOU TO TRY TO RESPOND AS YOU WOULD IF THESE WERE ACTUAL SITUATIONS. CONSIDER CAREFULLY WHAT EACH GRADE POINT AVERAGE WOULD MEAN TO YOU AT GRADUATION AND HOW EACH WOULD AFFECT YOUR GOALS FOR FURTHER EDUCATION AND FUTURE EMPLOYMENT.

THESE GAMBLING CANNOT, OF COURSE, BE CARRIED OUT IN REAL LIFE; HOWEVER, THE SITUATION SHOULD BE MEANINGFUL TO YOU. THE PURPOSE OF THE EXPERIMENT IS TO STUDY GAMBLING BEHAVIOR WHEN THE OUTCOMES HAVE SUBSTANTIAL IMPORTANCE RELATING TO A STUDENT'S FUTURE.

**experiment 1B**

THE PURPOSE OF THIS EXPERIMENT IS TO STUDY GAMBLING BEHAVIOR WHEN THE OBJECT OF THE GAMBLE IS AN IMPORTANT FACTOR RELATING TO THE QUALITY OF THE GRADUATE STUDENTS IN A PHD PROGRAM, NAMELY, THE AVERAGE OF THEIR VERBAL AND QUANTITATIVE GRADUATE RECORD EXAMINATION (GRE) SCORES. YOU ARE TO THINK OF YOURSELF AS THE CHAIR OF THE ADMISSIONS COMMITTEE OF THE GRADUATE DEPARTMENT WITH WHICH YOU CAN MOST COMFORTABLY IDENTIFY.

AS CHAIR OF THE COMMITTEE, YOUR OPINIONS ARE HIGHLY REGARDED AND THE STUDENTS YOU RECOMMEND ARE LIKELY TO BE SENT ACCEPTANCE LETTERS. YOU, OF COURSE, WOULD LIKE TO ADMIT THE STUDENTS WITH THE HIGHEST AVERAGE GRE SCORES, GIVEN THAT ALL OTHER RATING FACTORS ARE EQUAL.

FOR THE PURPOSES OF THIS EXPERIMENT, HOWEVER, A MORE DIFFICULT TASK WILL BE PRESENTED TO YOU. YOU CANNOT SIMPLY RECOMMEND THE STUDENTS WITH THE TOP TEST SCORES. RATHER, YOU WILL BE PRESENTED WITH SEVERAL SITUATIONS INVOLVING CHOICES BETWEEN GETTING A STUDENT WITH A PARTICULAR AVERAGE GRE SCORE FOR SURE, AND A GAMBLE WHICH WILL GET YOU A STUDENT WITH A HIGHER SCORE IF YOU WIN OR ONE WITH A LOWER SCORE IF YOU LOSE.

WHILE WE REALIZE THAT YOU ARE NOT ACCUSTOMED TO GAMBLING ON THE GRE SCORES OF THE STUDENTS TO BE ADMITTED TO YOUR GRADUATE PROGRAM, WE WOULD LIKE YOU TO RESPOND AS YOU WOULD IF THESE WERE ACTUAL DECISION SITUATIONS. CONSIDER CAREFULLY WHAT IT WOULD MEAN TO YOU, YOUR UNIVERSITY, AND ALL OTHERS CONCERNED TO ADMIT A GRADUATE STUDENT WITH A PARTICULAR AVERAGE GRE SCORE.

THE SITUATIONS THAT WILL BE PRESENTED TO YOU CANNOT, OF COURSE, BE CARRIED OUT IN REAL LIFE; HOWEVER, YOUR RESPONSES TO THESE IDEALIZED SITUATIONS WILL TELL US SOMETHING ABOUT YOUR BEHAVIOR IN MORE REALISTIC SITUATIONS. THE PURPOSE OF THE EXPERIMENT IS TO STUDY GAMBLING BEHAVIOR WHEN THE OUTCOMES HAVE SUBSTANTIAL IMPORTANCE RELATING TO THE NORMAL FUNCTIONING OF A GRADUATE INSTITUTION.

#### B. General Instructions - Both Formats

IN THIS EXPERIMENT YOU WILL BE PRESENTED WITH SEVERAL HYPOTHETICAL CHOICE SITUATIONS CONSISTING OF A FOR-SURE AND A CHANCE OPTION. THE FOR-SURE OPTION OFFERS YOU THE CERTAINTY OF KNOWING THE OUTCOME IF YOU CHOOSE THAT OPTION. THE CHANCE OPTION OFFERS YOU A PROBABILISTIC CHANCE AT TWO POSSIBLE OUTCOMES.

FOR EXAMPLE, THE FOR-SURE OPTION COULD HAVE YOU ADMIT A STUDENT WITH AN AVERAGE GRE SCORE OF 550. THE CHANCE OPTION COULD THEN OFFER YOU THE CHANCE OF ADMITTING SOMEONE WITH A SCORE OF 750 WITH PROBABILITY  $P$  (WHERE  $P$  CAN BE ANY VALUE BETWEEN ZERO AND ONE) OR A PERSON WITH A SCORE OF 500 WITH PROBABILITY  $1-P$ .

ASSUMING THAT YOU PREFER TO ADMIT A STUDENT WITH A SCORE OF 750 TO ONE WITH A SCORE OF 550 TO ONE WITH A SCORE OF 500 AND ASSUMING THAT YOU ARE GIVEN A CHOICE BETWEEN THE FOR-SURE AND CHANCE OPTIONS, YOUR DECISION AS TO WHICH OPTION YOU WOULD PREFER TO TAKE WILL DEPEND UPON THE VALUE OF  $P$ , THE PROBABILITY OF ADMITTING A STUDENT WITH AN AVERAGE GRE SCORE OF 750 UNDER THE CHANCE OPTION.

THIS PARTICULAR CHOICE OF ADMITTING A STUDENT WITH A SCORE OF 550 FOR SURE VERSUS THE CHANCE OF ADMITTING STUDENTS WITH SCORES OF 750 AND 500 WITH PROBABILITIES  $P$  AND  $1-P$ , RESPECTIVELY, WOULD BE DISPLAYED ON THE SCREEN AS FOLLOWS.

I		GRE	I	I		GRE		I
I			I	I		750	P	I
I	FOR SURE	550	I	I	CHANCE			I
I			I	I		500	1-P	I
I			I	I				I

FOR P VALUES CLOSE TO ONE YOU WILL PREFER THE CHANCE OPTION BECAUSE IT WILL PROBABLY RESULT IN THE ADMITTANCE OF A STUDENT WITH AN AVERAGE GRE SCORE OF 750 RATHER THAN THE ADMITTANCE OF ONE WITH A SCORE OF 500. THE FOR-SURE OPTION, ON THE OTHER HAND, SAYS THAT THE ADMITTED STUDENT WILL ONLY HAVE A SCORE OF 550.

FOR VALUES NEAR ZERO, HOWEVER, YOU WILL PREFER THE SURE THING (ADMITTING A STUDENT WITH A SCORE OF 550) BECAUSE THE CHANCE OPTION WILL PROBABLY RESULT IN THE ADMITTANCE OF A STUDENT WITH A GRE SCORE OF 500. THIS IS BECAUSE A P VALUE NEAR ZERO MEANS THAT THE VALUE OF  $1-P$ , THE PROBABILITY OF ADMITTING SOMEONE WITH AN AVERAGE GRE SCORE OF 500, WILL BE NEAR ONE.

FOR SOME UNIQUE INTERMEDIATE VALUE OF P YOU SHOULD BE INDIFFERENT BETWEEN THE CHANCE AND THE FOR-SURE OPTIONS. THIS MEANS THAT WHEN P IS AT THIS VALUE YOU WOULD BE CONTENT TO HAVE THE OPTION, EITHER FOR-SURE OR CHANCE, YOU RECEIVE DETERMINED BY CHANCE.

WE ARE INTERESTED IN THIS VALUE OF P, THE INDIFFERENCE PROBABILITY, FOR EACH SITUATION (SET OF CHANCE AND FOR-SURE OPTIONS) WITH WHICH YOU WILL BE PRESENTED.

ALL OF THE SITUATIONS WITH WHICH YOU WILL BE PRESENTED WILL INVOLVE AVERAGE GRE SCORES AND THESE SCORES ARE MEANT TO REFER TO THE TEST SCORES OF PROSPECTIVE GRADUATE STUDENTS IN THE DEPARTMENT FOR WHICH, IN THIS EXPERIMENT, YOU ARE SERVING AS CHAIR OF THE ADMISSIONS COMMITTEE. PLEASE THINK CAREFULLY HOW THE ADMITTANCE OF A STUDENT WITH A PARTICULAR AVERAGE GRE SCORE WILL AFFECT YOURSELF, YOUR DEPARTMENT, AND ALL OTHERS CONCERNED.

### C. Specialized Instructions - Format One

IN THIS HALF OF THE EXPERIMENT THE COMPUTER WILL GENERATE THE SITUATIONS AND YOU ARE TO TELL THE EXPERIMENTER THE VALUE OF P THAT WILL MAKE YOU INDIFFERENT BETWEEN THE FOR-SURE AND CHANCE OPTIONS. REMEMBER, WHEN THE PROBABILITY P OF ADMITTING THE STUDENT WITH THE HIGHER AVERAGE GRE SCORE IN THE CHANCE OPTION IS AT YOUR INDIFFERENCE PROBABILITY, YOU WILL LIKE THE CHANCE AND THE FOR-SURE OPTIONS EQUALLY WELL.

WE REALIZE THAT YOUR INDIFFERENCE PROBABILITIES WILL PROBABLY BE SOMEWHAT VAGUE. SINCE THERE ARE NO RIGHT OR WRONG ANSWERS, JUST TRY TO RESPOND AS DELIBERATELY, CAREFULLY AND ACCURATELY AS YOU CAN. HOWEVER, IF YOU MAKE A MISTAKE OR CHANGE YOUR MIND, DON'T WORRY, YOU WILL HAVE AN OPPORTUNITY TO CHANGE YOUR RESPONSES LATER.

### **Specialized Instructions - Format Two**

IN THIS HALF OF THE EXPERIMENT, THE COMPUTER WILL GENERATE A POSSIBLE VALUE FOR P AND YOU ARE TO INDICATE TO THE RESEARCH ASSISTANT WHETHER THIS PROBABILITY WILL MAKE YOU INDIFFERENT BETWEEN THE FOR-SURE AND CHANCE OPTIONS. REMEMBER, IF YOU ARE INDIFFERENT BETWEEN THE TWO OPTIONS YOU WOULD NOT PREFER TO TAKE ONE OPTION OVER THE OTHER.

IF THE P VALUE PRESENTED DOES NOT MAKE YOU INDIFFERENT BETWEEN THE OPTIONS, YOU ARE TO INDICATE WHICH OPTION, EITHER FOR-SURE OR CHANCE, YOU WOULD PREFER TO TAKE. THE COMPUTER WILL CONTINUE TO GENERATE P VALUES UNTIL YOU INDICATE THAT THE P VALUE PRESENTED MAKES YOU INDIFFERENT BETWEEN THE FOR-SURE AND CHANCE OPTIONS.

ALTHOUGH YOUR INDIFFERENCE PROBABILITIES WILL UNDOUBTEDLY BE SOMEWHAT VAGUE, TRY TO ANSWER AS ACCURATELY AS POSSIBLE. THERE ARE NO RIGHT OR WRONG ANSWERS. IF YOU MAKE A MISTAKE OR CHANGE YOUR MIND AS TO WHICH OPTION YOU PREFER, INFORM THE EXPERIMENTER AND SHE WILL MAKE THE APPROPRIATE CHANGE. IF YOU CHANGE YOUR MIND AFTER STATING THAT YOU ARE INDIFFERENT, YOU MAY CHANGE THE INDIFFERENCE PROBABILITY LATER, AFTER YOU HAVE FINISHED ALL OF THE SITUATIONS.

#### **D. Before Seeing the Fitted Probabilities**

SPECIFYING INDIFFERENCE PROBABILITIES IS A DIFFICULT TASK, AND MOST PEOPLE, INCLUDING TRAINED STATISTICIANS, ARE UNABLE TO PROVIDE A SET OF PROBABILITIES THAT ARE CONSISTENT WITH EACH OTHER ON THE FIRST TRY. USING PROBABILITIES THAT ARE INCONSISTENT WILL USUALLY RESULT IN THE ADMITTANCE OF STUDENTS WITH A LOWER AVERAGE GRE SCORE THAN WOULD BE ADMITTED IF THE PROBABILITIES WERE CONSISTENT.

SINCE THE AVERAGE GRE SCORES OF THE STUDENTS YOU ADMIT WILL HAVE SUBSTANTIAL EFFECT ON THE REPUTATION OF YOUR DEPARTMENT, IT IS VERY IMPORTANT TO MAKE SURE THAT THE INDIFFERENCE PROBABILITIES YOU SPECIFY ARE CONSISTENT. WITH A LITTLE HELP THIS IS POSSIBLE.

IN ORDER TO HELP YOU, THE COMPUTER HAS BEEN PROGRAMMED TO PROVIDE A SET OF INTERNALLY CONSISTENT PROBABILITIES THAT IS SIMILAR TO THE INCONSISTENT SET YOU SPECIFIED. SUCH AN INTERNALLY CONSISTENT SET OF PROBABILITIES IS SAID TO BE COHERENT.

THE COMPUTER HAS FINISHED CALCULATING THE COHERENT SET OF INDIFFERENCE PROBABILITIES. EXAMINE THEM CAREFULLY AND COMPARE THEM TO THE ONES YOU SPECIFIED. IF YOU FEEL THAT THE NEW SET OF PROBABILITIES ACCURATELY REFLECTS YOUR BETTING BEHAVIOR YOU MAY ACCEPT THOSE PROBABILITIES.

SINCE THE COMPUTER BASES ITS RECOMMENDED PROBABILITIES SOLELY UPON THE INDIFFERENCE PROBABILITIES SPECIFIED AND HAS NO REAL IDEA OF THE TRUE FEELINGS ABOUT GRE SCORES UNDERLYING THEM, THE PROBABILITIES MAY NOT ACCURATELY REFLECT THOSE FEELINGS.

IF YOU FEEL THAT THIS IS THE CASE YOU MAY CHANGE ANY OR ALL OF YOUR SPECIFIED PROBABILITIES. THE COMPUTER WILL THEN GENERATE A NEW SET OF COHERENT PROBABILITIES BASED ON YOUR REVISED ESTIMATES.

THE FOLLOWING TABLE SUMMARIZES THE GAMBLER, ALONG WITH THE SPECIFIED AND COMPUTER-GENERATED INDIFFERENCE PROBABILITIES. COMPARE THE TWO SETS OF PROBABILITIES AND DECIDE WHETHER OR NOT THE COMPUTER-GENERATED SET ACCURATELY DESCRIBES YOUR ATTITUDES TOWARD THE VARIOUS TEST SCORES.

IF SO, YOU MAY DECIDE TO ACCEPT THE NEW SET OF PROBABILITIES. IF NOT, YOU MAY CHANGE SOME OR ALL OF THE ORIGINAL INDIFFERENCE PROBABILITIES YOU SPECIFIED. THINK CAREFULLY BEFORE YOU MAKE YOUR DECISION.



## PILOT EXPERIMENT 2

The instructions for this experiment were basically the same as those for Experiments 1 and 3. The instructions for using Format Four (modified ends-in) may be found on page 44. The Format Three (quartering method) explanation is reprinted below.

IN THIS HALF OF THE EXPERIMENT YOU WILL BE PRESENTED WITH THE FOLLOWING KIND OF TABLE, REPRESENTING ONE SITUATION, AND YOU ARE TO INDICATE WHICH OPTION, EITHER FOR-SURE OR CHANCE, YOU WOULD PREFER TO TAKE:

FOR SURE OPTION			CHANCE OPTION			
I	GPA	I	I	GPA	CHANCE	I
I		I	I	2.5	85%	I
I	0.5 FOR SURE	I	I			I
I		I	I	0.0	15%	I

ALTHOUGH YOUR RESPONSES WILL PROBABLY BE SOMEWHAT VAGUE, TRY TO ANSWER AS ACCURATELY AS YOU CAN. THERE ARE NO RIGHT OR WRONG ANSWERS. IF YOU MAKE A MISTAKE OR CHANGE YOUR MIND AS TO WHICH OPTION YOU PREFER, INFORM THE RESEARCH ASSISTANT AND SHE WILL MAKE THE APPROPRIATE CHANGE.

### EXPERIMENT 3

#### A. Introduction - SPS, RC, LC

THE PURPOSE OF THIS EXPERIMENT IS TO STUDY GAMBLING BEHAVIOR WHEN THE OBJECT OF THE GAMBLE IS AN IMPORTANT FACTOR RELATING TO A COLLEGE STUDENT'S FUTURE OPPORTUNITIES, NAMELY, CUMULATIVE GRADE POINT AVERAGE (GPA) AT GRADUATION.

YOU WILL BE PRESENTED WITH SEVERAL SITUATIONS INVOLVING CHOICES BETWEEN GAMBLES AND SURE THINGS. YOU ARE TO IMAGINE THAT THE OUTCOME OF THE OPTION YOU CHOOSE WILL RESULT IN THE FIXING OF A PARTICULAR GPA ON YOUR PERMANENT UNIVERSITY RECORD.

WHILE WE REALIZE THAT YOU ARE NOT ACCUSTOMED TO GAMBLING ON YOUR GRADE POINT AVERAGE, WE WOULD LIKE YOU TO TRY TO RESPOND AS YOU WOULD IF THESE WERE ACTUAL SITUATIONS. CONSIDER CAREFULLY WHAT EACH GRADE POINT AVERAGE WOULD MEAN TO YOU AT GRADUATION AND HOW EACH WOULD AFFECT YOUR GOALS FOR FURTHER EDUCATION AND FUTURE EMPLOYMENT.

THESE GAMBLES CANNOT, OF COURSE, BE CARRIED OUT IN REAL LIFE; HOWEVER, THE SITUATIONS SHOULD BE MEANINGFUL TO YOU. THE PURPOSE OF THE EXPERIMENT IS TO STUDY GAMBLING BEHAVIOR WHEN THE OUTCOMES HAVE SUBSTANTIAL IMPORTANCE RELATING TO A STUDENT'S FUTURE.

#### B. General Instructions - SPS, LC

IN THIS EXPERIMENT YOU WILL BE PRESENTED WITH SEVERAL HYPOTHETICAL CHOICE SITUATIONS CONSISTING OF A FOR-SURE AND A CHANCE OPTION. THE FOR-SURE OPTION OFFERS YOU THE CERTAINTY OF KNOWING THE OUTCOME IF YOU CHOOSE THAT OPTION. THE CHANCE OPTION OFFERS YOU A PROBABILISTIC CHANCE AT TWO POSSIBLE OUTCOMES.

FOR EXAMPLE, THE FOR-SURE OPTION COULD OFFER YOU A GPA OF 0.5. THE CHANCE OPTION COULD THEN OFFER YOU A CHANCE OF GETTING A GPA OF 2.5 WITH PROBABILITY  $P$  (WHERE  $P$  CAN BE ANY VALUE BETWEEN 5% AND 95%) AND A CHANCE ON GETTING A GPA OF 0.0 WITH PROBABILITY  $100\% - P$ .

ASSUMING THAT YOUR PREFERENCE FOR GPA'S IS 2.5 OVER 0.5 OVER 0.0 AND ASSUMING THAT YOU ARE GIVEN A CHOICE BETWEEN THE FOR-SURE AND THE CHANCE OPTIONS, YOUR DECISION AS TO WHICH OPTION YOU WOULD PREFER TO TAKE WILL DEPEND UPON THE VALUE OF  $P$ , THE PROBABILITY OF GETTING A GPA OF 2.5 UNDER THE CHANCE OPTION.

THE CHANCE AND FOR-SURE OPTIONS WILL BE PRESENTED TO YOU IN SEPARATE BOXES ON THE SCREEN AND THE PROBABILITIES ASSOCIATED WITH EACH OF THE POSSIBLE GPAS IN THE CHANCE OPTION WILL BE CLEARLY MARKED.

A LITTLE THOUGHT SHOULD CONVINCE YOU THAT THE STATEMENTS IN THE FOLLOWING THREE PARAGRAPHS ARE TRUE. HOWEVER, THE CONCEPTS MAY SEEM SOMEWHAT DIFFICULT SO IF YOU HAVE ANY QUESTIONS OR ARE AT ALL UNSURE OF WHAT IS BEING STATED, PLEASE DO NOT HESITATE TO ASK THE RESEARCH ASSISTANT TO EXPLAIN THINGS MORE CLEARLY. SHE WILL BE GLAD TO CLARIFY ANYTHING YOU DO NOT QUITE UNDERSTAND.

FOR P VALUES CLOSE TO 100% YOU WILL PREFER THE CHANCE OPTION BECAUSE IT WILL PROBABLY YIELD YOU A GRADE POINT AVERAGE OF 2.5 RATHER THAN A GPA OF 0.0. THE FOR-SURE OPTION, ON THE OTHER HAND, WILL ONLY NET YOU A GPA OF 0.5.

FOR VALUES NEAR 0%, HOWEVER, YOU WILL PREFER THE SURE THING (A GPA OF 0.5) BECAUSE THE CHANCE OPTION WILL PROBABLY YIELD YOU A GRADE POINT OF 0.0. THIS IS BECAUSE A P VALUE CLOSE TO 0% MEANS THAT THE VALUE OF  $100\% - P$ , THE PROBABILITY OF GETTING A GPA OF 0.0, WILL BE NEAR 100%.

FOR SOME UNIQUE INTERMEDIATE VALUE OF P YOU SHOULD BE INDIFFERENT BETWEEN THE CHANCE AND THE FOR-SURE OPTIONS. THIS MEANS THAT WHEN P IS AT THIS VALUE YOU WOULD BE CONTENT TO LET SOMEONE ELSE DETERMINE WHICH OPTION, EITHER FOR-SURE OR CHANCE, YOU RECEIVE.

WE ARE INTERESTED IN THIS VALUE OF P, THE INDIFFERENCE PROBABILITY, FOR EACH SITUATION (SET OF CHANCE AND FOR-SURE OPTIONS) WITH WHICH YOU WILL BE PRESENTED.

ALL OF THE SITUATIONS WITH WHICH YOU WILL BE PRESENTED WILL INVOLVE GPA'S, AND THE FOLLOWING SCALE FOR GRADE POINTS WILL BE USED:

4.0=A  
3.0=B  
2.0=C  
1.0=D  
0.0=F

THESE GPA'S ARE MEANT TO REFER TO YOUR GPA UPON GRADUATION FROM YOUR PRESENT UNIVERSITY. IN RESPONDING TO THE CHOICES THAT WILL BE PRESENTED TO YOU, PLEASE THINK CAREFULLY ABOUT HOW THE ENTRY OF ANY PARTICULAR GPA ON YOUR PERMANENT RECORD WILL AFFECT YOUR CAREER AND LIFE.

General Instructions - RC

IN THIS EXPERIMENT YOU WILL BE PRESENTED WITH SEVERAL HYPOTHETICAL CHOICE SITUATIONS CONSISTING OF A FOR-SURE AND A CHANCE OPTION. THE FOR-SURE OPTION OFFERS YOU THE CERTAINTY OF KNOWING THE OUTCOME IF YOU CHOOSE THAT OPTION. THE CHANCE OPTION OFFERS YOU A PROBABILISTIC CHANCE AT TWO POSSIBLE OUTCOMES.

FOR EXAMPLE, THE FOR-SURE OPTION COULD OFFER YOU A GPA OF 2.5. THE CHANCE OPTION COULD THEN OFFER YOU A CHANCE OF GETTING A GPA OF 3.5 WITH PROBABILITY  $P$  (WHERE  $P$  CAN BE ANY VALUE BETWEEN 5% AND 95%) AND A CHANCE OF GETTING A GPA OF 1.5 WITH PROBABILITY  $100\% - P$ .

ASSUMING THAT YOUR PREFERENCE FOR GPA'S IS 3.5 OVER 2.5 OVER 1.5 AND ASSUMING THAT YOU ARE GIVEN A CHOICE BETWEEN THE FOR-SURE AND THE CHANCE OPTIONS, YOUR DECISION AS TO WHICH OPTION YOU WOULD PREFER TO TAKE WILL DEPEND UPON THE VALUE OF  $P$ , THE PROBABILITY OF GETTING A GPA OF 3.5 UNDER THE CHANCE OPTION.

THE CHANCE AND FOR-SURE OPTIONS WILL BE PRESENTED TO YOU IN BOXES ON THE SCREEN, AND THE PROBABILITIES ASSOCIATED WITH EACH OF THE POSSIBLE GPA'S IN THE CHANCE OPTION WILL BE CLEARLY MARKED.

A LITTLE THOUGHT SHOULD CONVINCE YOU THAT THE STATEMENTS IN THE FOLLOWING THREE PARAGRAPHS ARE TRUE. HOWEVER, IF YOU HAVE ANY QUESTIONS OR ARE AT ALL UNSURE OF WHAT IS BEING STATED, PLEASE DO NOT HESITATE TO ASK THE RESEARCH ASSISTANT TO EXPLAIN THINGS MORE CLEARLY. SHE WILL BE GLAD TO CLARIFY ANYTHING THAT YOU DO NOT QUITE UNDERSTAND.

FOR  $P$  VALUES CLOSE TO 100% YOU WILL PREFER THE CHANCE OPTION BECAUSE IT WILL PROBABLY YIELD YOU A GRADE POINT AVERAGE OF 3.5 RATHER THAN A GPA OF 1.5. THE FOR-SURE OPTION, ON THE OTHER HAND, WILL ONLY NET YOU A GPA OF 2.5.

FOR VALUES NEAR 0%, HOWEVER, YOU WILL PREFER THE SURE THING (A GPA OF 2.5) BECAUSE THE CHANCE OPTION WILL PROBABLY YIELD YOU A GRADE POINT OF 1.5. THIS IS BECAUSE A  $P$  VALUE CLOSE TO 0% MEANS THAT THE VALUE OF  $100\% - P$ , THE PROBABILITY OF GETTING A GPA OF 1.5, WILL BE NEAR 100%.

FOR SOME UNIQUE INTERMEDIATE VALUE OF  $P$  YOU SHOULD BE INDIFFERENT BETWEEN THE CHANCE AND THE FOR-SURE OPTIONS. THIS MEANS THAT WHEN  $P$  IS AT THIS VALUE YOU WOULD BE CONTENT TO LET SOMEONE ELSE DETERMINE WHICH OPTION, EITHER FOR-SURE OR CHANCE, YOU RECEIVE.

WE ARE INTERESTED IN THIS VALUE OF  $P$ , THE INDIFFERENCE PROBABILITY, FOR EACH SITUATION (SET OF CHANCE AND FOR-SURE OPTIONS) WITH WHICH YOU WILL BE PRESENTED.

ALL OF THE SITUATIONS WITH WHICH YOU WILL BE PRESENTED WILL INVOLVE GPA'S, AND THE FOLLOWING SCALE FOR GRADE POINTS WILL BE USED:

4.0=A  
3.0=B  
2.0=C  
1.0=D  
0.0=F

THESE GPA'S ARE MEANT TO REFER TO YOUR GPA UPON GRADUATION FROM YOUR PRESENT UNIVERSITY. IN RESPONDING TO THE CHOICES THAT WILL BE PRESENTED TO YOU, PLEASE THINK CAREFULLY ABOUT HOW THE ENTRY OF ANY PARTICULAR GPA ON YOUR PERMANENT RECORD WILL AFFECT YOUR CAREER AND LIFE.

### C. Procedure Specific Instructions - RC

TO HELP YOU DETERMINE WHETHER THE INDIFFERENCE PROBABILITIES YOU HAVE SPECIFIED ARE CONSISTENT WITH EACH OTHER, WE HAVE DEVISED THE FOLLOWING PROCEDURE: AFTER YOU HAVE CHOSEN TWO INDIFFERENCE PROBABILITIES, THE COMPUTER WILL PRESENT YOU WITH A TABLE WHICH DISPLAYS THESE TWO SITUATIONS AND ALSO TWO MORE SITUATIONS FOR WHICH IT HAS DEDUCED PROBABILITIES BASED ON THE TWO YOU SPECIFIED.

THINK AGAIN OF THE SITUATION IN WHICH YOU WERE OFFERED A 'FOR-SURE' GPA OF 2.5 AND A CHANCE OF GETTING A GPA OF 3.5 WITH PROBABILITY P AND OF GETTING A GPA OF 1.5 WITH PROBABILITY 100%-P. LET US IMAGINE THAT YOU SPECIFIED P TO BE 60%. NOW THINK OF A SITUATION IN WHICH YOU ARE OFFERED A 'FOR-SURE' GPA OF 3.0 VERSUS A CHANCE OF GETTING A GPA OF 3.5 WITH PROBABILITY P AND OF GETTING A GPA OF 2.5 WITH PROBABILITY 100%-P. LET US IMAGINE THAT FOR THIS SITUATION YOU CHOOSE A P VALUE OF 65%.

	SITUATIONS			
	1	2	3	4
P CHANCE	3.5	3.5	3.0	3.5
FOR SURE	2.5	3.0	2.5	3.0
100%-P CHANCE	1.5	2.5	1.5	1.5
	P=60%	P=65%	P=70%	P=86%

YOU WILL NOW BE INVITED TO EXAMINE ALL FOUR SITUATIONS AND TO DETERMINE WHETHER THE INDIFFERENCE PROBABILITIES DISPLAYED FOR EACH ARE CONSISTENT WITH YOUR VIEWS. IF YOU RECOGNIZE SOME INCONSISTENCIES YOU WILL BE ABLE TO ADJUST THE P VALUES FOR ANY TWO OF THE SITUATIONS (THEREBY FIXING THE P VALUES FOR THE OTHER TWO SITUATIONS). YOU WILL BE ASKED TO MANIPULATE THE P VALUES UNTIL YOU ARE REASONABLY SATISFIED WITH THE DISPLAYED INDIFFERENCE PROBABILITIES FOR ALL FOUR SITUATIONS.

IT MAY TAKE SOME TIME TO FIND THE P VALUES THAT WILL MAKE YOU SATISFIED IN ALL OF THE SITUATIONS, SO WE HOPE THAT YOU WILL BE PATIENT AND PERSEVERING.

WE REALIZE THAT SPECIFYING INDIFFERENCE PROBABILITIES FOR ISOLATED SITUATIONS IS A DIFFICULT TASK -- EVEN TRAINED STATISTICIANS FIND IT DIFFICULT. BECAUSE IT IS HARD TO SEE HOW THE VARIOUS SITUATIONS RELATE TO EACH OTHER, THE INDIFFERENCE PROBABILITIES FOR THEM ARE LIKELY TO BE INCONSISTENT WITH EACH OTHER. IF YOUR FINAL GPA WERE ACTUALLY DETERMINED BY HAVING YOU PLAY THROUGH SEVERAL SITUATIONS, IT WOULD BE HIGHER IF THE INDIFFERENCE PROBABILITIES FOR THE VARIOUS SITUATIONS WERE CONSISTENT WITH EACH OTHER THAN IF THEY WERE INCONSISTENT. A SET OF CONSISTENT PROBABILITIES IS SAID TO BE COHERENT.

SINCE YOUR FINAL GPA IS LIKELY TO AFFECT YOUR FUTURE OPPORTUNITIES, IT IS IMPORTANT THAT YOU TRY TO MAKE THE INDIFFERENCE PROBABILITIES AS CONSISTENT WITH EACH OTHER AS YOU CAN. THE PROCEDURE EXPLAINED ABOVE WAS DESIGNED TO HELP YOU IN THIS TASK. IF YOU ARE INDIFFERENT BETWEEN THE TWO OPTIONS PRESENTED IN EACH OF THE FOUR SITUATIONS FOR THE STATED PROBABILITIES, THEN ALL OF THOSE INDIFFERENCE PROBABILITIES WILL BE CONSISTENT WITH EACH OTHER.

#### Procedure Specific Instructions - LC

TO HELP YOU DETERMINE WHETHER THE INDIFFERENCE PROBABILITY YOU HAVE CHOSEN ACCURATELY REFLECTS YOUR VIEWS, WE HAVE PROVIDED A METHOD BY WHICH YOU CAN TEST YOUR INITIAL JUDGMENT AGAINST A JUDGMENT IN A RELATED SITUATION.

THINK AGAIN ABOUT THE SITUATION IN WHICH YOU WERE OFFERED A 'FOR-SURE' GPA OF 0.5 VERSUS A CHANCE OF GETTING A GPA OF 2.5 WITH PROBABILITY P AND OF GETTING A GPA OF 0.0 WITH PROBABILITY 100%-P. LET US IMAGINE THAT YOU SPECIFIED P TO BE 50%.

THE COMPUTER PROGRAM THEN REPRODUCES THIS SITUATION (WITH P=50%) AND ALSO PRESENTS YOU WITH ANOTHER TYPE OF CHOICE SITUATION CONSTRUCTED OF TWO CHANCE OPTIONS, WITH THE PROBABILITIES FOR THE NEW SITUATION DEDUCED FROM THE INDIFFERENCE PROBABILITY YOU SPECIFIED FOR THE FIRST SITUATION.

#### HERE ARE THE TWO SITUATIONS

##### GAMBLE SITUATION 1

I	OPTION 1	OPTION 2	I
I	(FOR SURE)	(CHANCE)	I
I		2.5 50%	I
I	0.5		I
I		0.0 50%	I

##### GAMBLE SITUATION 2

I	OPTION 1	OPTION 2	I
I			I
I	50%	2.5 75%	I
I	50%	0.5 --	I
I	--	0.0 25%	I

RECALL THAT SITUATION 1 SAYS THAT YOU HAVE NO PREFERENCE BETWEEN GETTING A GPA OF 0.5 FOR SURE AND TAKING A CHANCE OPTION WHICH GIVES YOU A 50% CHANCE OF GETTING A GPA OF 2.5 AND A 50% CHANCE OF GETTING A GPA OF 0.0.

YOU ARE NOW ASKED TO DETERMINE WHETHER OR NOT YOU ARE INDIFFERENT BETWEEN THE TWO OPTIONS IN SITUATION 2. NOTE THAT OPTION 2 GIVES YOU A BETTER CHANCE AT THE MOST PREFERRED OUTCOME (2.5) BUT ALSO SOME POSSIBILITY OF GETTING THE LEAST PREFERRED OUTCOME (0.0).

THERE IS NO CHANCE OF GETTING THE MIDDLE GPA (0.5). OPTION 1, ON THE OTHER HAND, GIVES YOU A CHANCE OF GETTING THE MIDDLE GPA INSTEAD OF A CHANCE OF GETTING THE LOWEST GPA. THIS IMPROVEMENT (ASSUMING YOU PREFER TO HAVE A CHANCE OF GETTING THE 0.5 TO A CHANCE OF GETTING THE 0.0), HOWEVER, IS COMPENSATED BY THE LOWER PROBABILITY OF GETTING THE MOST PREFERRED GPA (A 50% CHANCE VERSUS A 75% CHANCE).

YOU WILL BE ASKED TO MANIPULATE THE CHANCE OF GETTING A GPA OF 2.5 IN SITUATION 1, OPTION 2 UNTIL A) YOU ARE INDIFFERENT BETWEEN OPTIONS 1 AND 2 IN SITUATION 2 AND B) YOU ARE ALSO INDIFFERENT BETWEEN THE FOR-SURE AND CHANCE OPTIONS IN SITUATION 1.

IT MAY TAKE SOME TIME BEFORE YOU ARE SATISFIED WITH (I.E., INDIFFERENT BETWEEN) BOTH SITUATIONS, SO WE HOPE THAT YOU WILL BE PATIENT AND PERSEVERING.

WE REALIZE THAT SPECIFYING INDIFFERENCE PROBABILITIES FOR ISOLATED SITUATIONS IS A DIFFICULT TASK -- EVEN TRAINED STATISTICIANS FIND IT DIFFICULT. BECAUSE IT IS HARD TO SEE HOW THE VARIOUS SITUATIONS RELATE TO EACH OTHER, THE INDIFFERENCE PROBABILITIES FOR THEM ARE LIKELY TO BE INCONSISTENT WITH EACH OTHER. IF YOUR FINAL GPA WERE ACTUALLY DETERMINED BY HAVING YOU PLAY THROUGH SEVERAL SITUATIONS, IT WOULD BE HIGHER IF INDIFFERENCE PROBABILITIES FOR THE VARIOUS SITUATIONS WERE CONSISTENT WITH EACH OTHER THAN IF THEY WERE INCONSISTENT. A SET OF CONSISTENT PROBABILITIES IS SAID TO BE COHERENT.

SINCE YOUR FINAL GPA IS LIKELY TO AFFECT YOUR FUTURE OPPORTUNITIES, IT IS IMPORTANT THAT YOU TRY TO MAKE THE INDIFFERENCE PROBABILITIES AS CONSISTENT WITH EACH OTHER AS YOU CAN. THE PROCEDURE EXPLAINED ABOVE WAS DESIGNED TO HELP YOU IN THIS TASK. IF YOU ARE INDIFFERENT BETWEEN THE OPTIONS PRESENTED IN BOTH SITUATIONS FOR THE STATED PROBABILITIES, THEN YOUR INDIFFERENCE PROBABILITIES WILL BE COHERENT.

#### D. Format Instructions - Direct Probability Elicitations (all procedures)

IN THIS HALF OF THE EXPERIMENT THE COMPUTER WILL GENERATE THE SITUATIONS AND YOU ARE TO TELL THE EXPERIMENTER THE VALUE OF P THAT WILL MAKE YOU INDIFFERENT BETWEEN THE FOR-SURE AND THE CHANCE OPTIONS.

HERE IS AN EXAMPLE:

FOR SURE OPTION				CHANCE OPTION			
I	GPA	I	I	I	GPA		I
I		I	I	I	2.5	PX	I
I	0.5	I	I	I			I
I		I	I	I	0.0	100%-P	I

WHAT VALUE OF P MAKES YOU INDIFFERENT BETWEEN THE FOR-SURE AND THE CHANCE OPTIONS?

REMEMBER, WHEN THE PROBABILITY P OF GETTING THE HIGHER GPA IN THE CHANCE OPTION IS AT YOUR INDIFFERENCE PROBABILITY, YOU WILL LIKE THE CHANCE AND THE FOR-SURE OPTIONS EQUALLY WELL.

WE REALIZE THAT YOUR INDIFFERENCE PROBABILITIES WILL UNDOUBTEDLY BE SOMEWHAT VAGUE. SINCE THERE ARE NO RIGHT OR WRONG ANSWERS, JUST TRY TO RESPOND AS DELIBERATELY, CAREFULLY AND ACCURATELY AS YOU CAN. HOWEVER, IF YOU MAKE A MISTAKE OR CHANGE YOUR MIND, DON'T WORRY, YOU WILL HAVE AN OPPORTUNITY TO CHANGE YOUR RESPONSES LATER.

**Format Instructions - Modified Ends-In (all procedures)**

IN THIS HALF OF THE EXPERIMENT YOU WILL BE PRESENTED WITH THE FOLLOWING KIND OF TABLE, REPRESENTING ONE SITUATION, AND YOU ARE TO INDICATE WHICH OPTION, EITHER FOR-SURE OR CHANCE, YOU WOULD PREFER TO TAKE:

FOR SURE OPTION				CHANCE OPTION			
I	GPA	I	I	I	GPA	I	I
I		I	I	I		I	I
I	0.5	I	I	I	0.0	I	2.5
I		I	I	I		I	I
100%				60%                      40%			

IF YOU HAVE NO PREFERENCE BETWEEN THE OPTIONS, STATE THAT YOU ARE INDIFFERENT. REMEMBER, IF YOU ARE INDIFFERENT BETWEEN THE TWO OPTIONS YOU WOULD CONSENT TO TO LET SOMEONE ELSE CHOOSE WHICH OPTION, EITHER FOR-SURE OR CHANCE, YOU RECEIVE. ALTHOUGH YOUR RESPONSES WILL UNDOUBTEDLY BE RATHER VAGUE, TRY TO ANSWER AS ACCURATELY AS YOU CAN. THERE ARE NO RIGHT OR WRONG ANSWERS. IF YOU MAKE A MISTAKE OR CHANGE YOUR MIND AS TO WHICH OPTION YOU PREFER, INFORM THE RESEARCH ASSISTANT AND SHE WILL MAKE THE APPROPRIATE CHANGE. YOUR FIRST SITUATION FOLLOWS.

**2. Procedure Specific Instructions - SFS**

SPECIFYING INDIFFERENCE PROBABILITIES IS A DIFFICULT TASK, AND MOST PEOPLE, INCLUDING TRAINED STATISTICIANS, ARE UNABLE TO PROVIDE A SET OF PROBABILITIES THAT ARE CONSISTENT WITH EACH OTHER ON THE FIRST TRY. USING PROBABILITIES THAT ARE INCONSISTENT WILL USUALLY RESULT IN THE APPEARANCE OF A LOWER FINAL GPA ON YOUR PERMANENT RECORD THAN WOULD APPEAR IF THE PROBABILITIES WERE CONSISTENT.

SINCE YOUR FINAL GRADE POINT AVERAGE WILL GREATLY AFFECT YOUR OPPORTUNITIES FOR THE FUTURE, IT IS VERY IMPORTANT TO MAKE SURE THAT THE INDIFFERENCE PROBABILITIES YOU SPECIFY ARE CONSISTENT. WITH A LITTLE HELP THIS IS POSSIBLE.

IN ORDER TO HELP YOU, THE COMPUTER HAS BEEN PROGRAMMED TO PROVIDE A SET OF INTERNALLY CONSISTENT PROBABILITIES THAT IS SIMILAR TO THE INCONSISTENT SET YOU SPECIFIED. SUCH AN INTERNALLY CONSISTENT SET OF PROBABILITIES IS SAID TO BE COHERENT.



SIT BACK AND RELAX FOR A FEW MINUTES WHILE THE COMPUTER MAKES THE NECESSARY CALCULATIONS.

THE COMPUTER HAS FINISHED CALCULATING THE COHERENT SET OF INDIFFERENCE PROBABILITIES. EXAMINE THEM CAREFULLY AND COMPARE THEM TO THE ONES YOU SPECIFIED. IF YOU FEEL THAT THE NEW SET OF PROBABILITIES ACCURATELY REFLECTS YOUR BETTING BEHAVIOR YOU MAY ACCEPT THOSE PROBABILITIES.

SINCE THE COMPUTER BASES ITS RECOMMENDED PROBABILITIES SOLELY UPON THE INDIFFERENCE PROBABILITIES SPECIFIED AND HAS NO REAL IDEA OF THE TRUE FEELINGS ABOUT GRADE POINT AVERAGES UNDERLYING THEM, THE PROBABILITIES MAY NOT ACCURATELY REFLECT THOSE FEELINGS.

IF YOU FEEL THAT THIS IS THE CASE YOU MAY CHANGE ANY OR ALL OF YOUR SPECIFIED PROBABILITIES. THE COMPUTER WILL THEN GENERATE A NEW SET OF COHERENT PROBABILITIES BASED ON YOUR REVISED ESTIMATES.

THE FOLLOWING TABLE SUMMARIZES THE GAMBLER'S, ALONG WITH THE SPECIFIED AND COMPUTER-GENERATED INDIFFERENCE PROBABILITIES. COMPARE THE TWO SETS OF PROBABILITIES AND DECIDE WHETHER OR NOT THE COMPUTER-GENERATED SET ACCURATELY DESCRIBES YOUR ATTITUDES TOWARD THE VARIOUS GRADE POINT AVERAGES.

IF SO, YOU MAY DECIDE TO ACCEPT THE NEW SET OF PROBABILITIES. IF NOT, YOU MAY CHANGE SOME OR ALL OF THE ORIGINAL INDIFFERENCE PROBABILITIES YOU SPECIFIED. THINK CAREFULLY BEFORE YOU MAKE YOUR DECISION.

SITUATION NO.	FOR SURE	CHANCE P	100%-P	INDIFF. P SPEC'D	COMP
1	0.50	1.00	0.00	50%	52%
2	1.00	2.00	0.00	50%	53%
3	1.50	2.50	0.50	50%	54%
4	2.00	2.50	1.50	50%	54%
5	2.50	4.00	1.00	65%	62%
6	3.00	3.50	2.50	55%	53%
7	3.50	4.00	3.00	70%	69%
8	1.00	1.50	0.50	50%	51%
9	1.50	2.00	1.00	50%	52%
10	1.50	3.00	0.00	60%	57%
11	2.00	3.00	1.00	50%	56%
12	2.00	4.00	0.00	70%	63%
13	2.50	3.00	2.00	50%	52%
14	2.50	3.50	1.50	55%	56%
15	3.00	4.00	2.00	55%	62%

- OPTIONS
1. MODIFY SPEC'D P'S
  2. ACCEPT COMP P'S

OPTION?

APPENDIX C - Questionnaire

1. Were you at any time unsure of what was expected of you?  
If so, when?
2. How could the directions have been changed so as to alleviate  
the above problem?
3. Are there any comments you would like to make about the  
experiment or the experimenter?
4. (Experiment 3 only) Please give some indication of how useful  
each GPA in the following scale is to you. The utility (usefulness  
or worth) of a GPA of 0.0 is assumed to be 0 while that of a GPA  
of 4.0 is 100. Please specify intermediate values for the other  
GPAs.

<u>GPA</u>	<u>utility to you</u>
0.0	0
0.5	
1.0	
1.5	
2.0	
2.5	
3.0	
3.5	
4.0	100

## Results Tables for Experiment 3

TABLE 4

Overall Mean Absolute Errors for Forty Subjects with Breakdowns  
by Procedure and Sex

	MALES			FEMALES		
	RC	SFS	LC	RC	SFS	LC
Format Order	.6574	.7736	.5976	.1311	.2317	.2405
	.2972	.3991	.3883	.1927	.0555	.4447
	.4372	.6445	.3433	.4662	.3237	.6073
					.3383	
4/1	.8540	.2099	.5727	.2037	.3917	.3550
	.4051	.4936	.2479	.3516	.3026	.5020
	.3105	.5067	.2180	.4409	.3342	.4059
	.4188				.6506	.6816
	$\bar{x} = .4803$ n=7	$\bar{x} = .5046$ n=6	$\bar{x} = .3938$ n=6	$\bar{x} = .2978$ n=6	$\bar{x} = .3285$ n=8	$\bar{x} = .4624$ n=7

Grand Mean (M) = .4065

Procedure:

$\bar{x}_{RC} = .3923$  (n = 13)  
 $\bar{x}_{SFS} = .4021$  (n = 14)  
 $\bar{x}_{LC} = .4295$  (n = 13)

Sex:

$\bar{x}_M = .4554$  (n = 19)  
 $\bar{x}_F = .3623$  (n = 21)

TABLE 5

Average Mean Absolute Errors for Forty Subjects Broken Down  
by Procedure, Format and Sex

FORMAT 1

	RC	SPS	LC	
Males	$\bar{x} = .5213$ n = 7	$\bar{x} = .4931$ n = 6	$\bar{x} = .3751$ n = 6	$\bar{x}_M = .4662$ n = 19
Females	$\bar{x} = .3085$ n = 6	$\bar{x} = .3112$ n = 8	$\bar{x} = .4589$ n = 7	$\bar{x}_F = .3597$ n = 21
	$\bar{x}_{RC} = .4231$ n = 13	$\bar{x}_{SPS} = .3892$ n = 14	$\bar{x}_{LC} = .4202$ n = 13	$\bar{x}_I = .4103$ n = 40

FORMAT 4

	RC	SPS	LC	
Males	$\bar{x} = .4275$ n = 7	$\bar{x} = .4974$ n = 6	$\bar{x} = .4122$ n = 6	$\bar{x}_M = .4447$ n = 19
Females	$\bar{x} = .2846$ n = 6	$\bar{x} = .3469$ n = 8	$\bar{x} = .4615$ n = 7	$\bar{x}_F = .3673$ n = 21
	$\bar{x}_{RC} = .3615$ n = 13	$\bar{x} = .4133$ n = 14	$\bar{x} = .4387$ n = 13	$\bar{x}_4 = .4040$ n = 40

TABLE 6

Differences in MAE (Format 1 MAE - Format 4 MAE)  
for Forty Subjects with Breakdowns by  
Procedure and Sex

	RC	SFS	LC
Males $\bar{x}_M = .0181$ n = 19	-.2174	.2592	.0490
	.0698	-.2262	.0262
	.0015	.2177	-.1275
	.0564	-.3650	.0305
	.0850	.1289	-.0408
	-.1339	-.0676	-.1599
	.7571		
	$\bar{x} = .0884$	$\bar{x} = -.0088$	$\bar{x} = -.0371$
	n = 7	n = 6	n = 6
Females $\bar{x}_F = -.0075$ n = 21	-.0079	-.0055	-.0169
	-.0887	-.0618	-.1135
	.1241	.0515	.1523
	.1431	.0282	.0049
	-.0372	.0049	.0135
	.0105	-.0302	-.1365
		-.1079	.0779
		-.1650	
	$\bar{x} = .0240$	$\bar{x} = -.0357$	$\bar{x} = -.0026$
	n = 6	n = 8	n = 7

M = .0046      N = 40

$\bar{x}_{RC} = .0587$   
n = 13

$\bar{x}_{SFS} = -.0242$   
n = 14

$\bar{x}_{LC} = -.0185$   
n = 13

TABLE 7

Greatest Absolute Differences in Utilities (GADU) for Forty Subjects with Breakdowns by Procedure, Format and Sex

Format  
1

MALES			FEMALES		
RC	LC	SFS	RC	LC	SFS
.2129	.1642	.6096	.4039	.1850	.5449
.1170	.3699	.3601	.1494	.4958	.5193
.1757	.2361	.5596	.5620	.4798	.1872
					.2343
$\bar{x} = .1685$ $n = 3$	$\bar{x} = .2569$ $n = 3$	$\bar{x} = .5098$ $n = 3$	$\bar{x} = .3718$ $n = 3$	$\bar{x} = .3869$ $n = 3$	$\bar{x} = .3714$ $n = 4$
.7205	.2477	.3635	.2844	.8104	.3758
.2763	.3210	.1178	.2070	.1759	.4199
.2145	.5064	.2234	.1380	.2826	.6599
.1153				.2940	.2165
$\bar{x} = .3316$ $n = 4$	$\bar{x} = .3584$ $n = 3$	$\bar{x} = .2349$ $n = 3$	$\bar{x} = .2098$ $n = 3$	$\bar{x} = .3908$ $n = 4$	$\bar{x} = .4167$ $n = 4$

Format  
4

Grand Mean (M) = .3384

Procedure:

$$\begin{aligned}\bar{x}_{RC} &= .2751 (n = 13) \\ \bar{x}_{LC} &= .3514 (n = 13) \\ \bar{x}_{SFS} &= .3848 (n = 14)\end{aligned}$$

Format:

$$\begin{aligned}\bar{x}_{F1} &= .3456 (n = 19) \\ \bar{x}_{F4} &= .3319 (n = 21)\end{aligned}$$

Sex:

$$\begin{aligned}\bar{x}_M &= .3111 (n = 19) \\ \bar{x}_F &= .3631 (n = 21)\end{aligned}$$

TABLE 8

## Two-way Summary Tables for Dependent Variable = GADU

## A. Summary Table for Procedure by Format

	RC	LC	SFS	
Format	$\bar{x} = .2701$	$\bar{x} = .3218$	$\bar{x} = .4307$	$\bar{x}_{F1} = .3456$
1	$n = 6$	$n = 6$	$n = 7$	$n = 19$
Format	$\bar{x} = .2794$	$\bar{x} = .3769$	$\bar{x} = .3388$	$\bar{x}_{F4} = .3319$
4	$n = 7$	$n = 7$	$n = 7$	$n = 21$
	$\bar{x}_{RC} = .2751$	$\bar{x}_{LC} = .3514$	$\bar{x}_{SFS} = .3848$	
	$n = 13$	$n = 13$	$n = 14$	

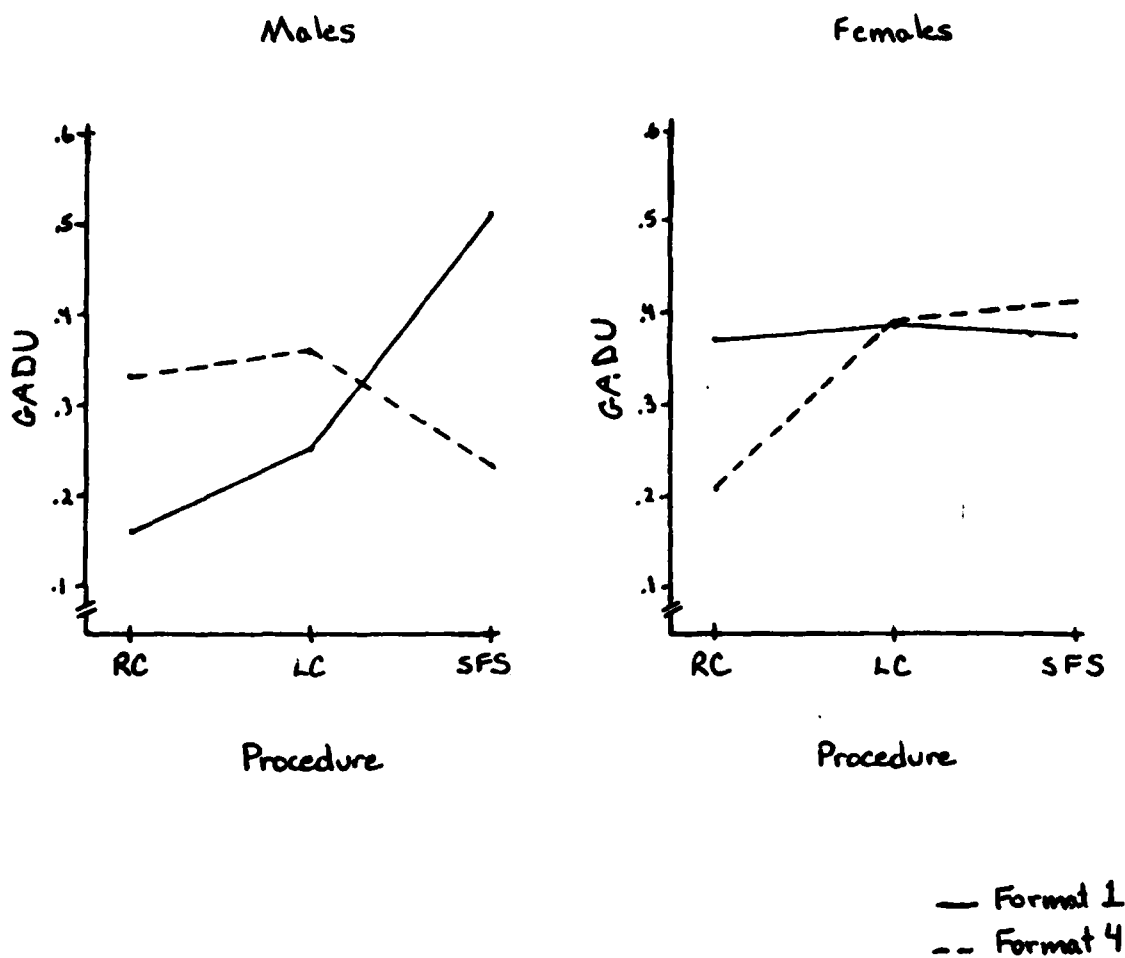
## B. Summary Table for Sex by Format

	MALES	FEMALES	
Format	$\bar{x} = .3117$	$\bar{x} = .3762$	$\bar{x}_{F1} = .3456$
1	$n = 9$	$n = 10$	$n = 19$
Format	$\bar{x} = .3106$	$\bar{x} = .3513$	$\bar{x}_{F4} = .3319$
4	$n = 10$	$n = 11$	$n = 21$
	$\bar{x}_M = .3111$	$\bar{x}_F = .3631$	
	$n = 19$	$n = 21$	

## C. Summary Table for Procedure by Sex

	RC	LC	SFS	
Males	$\bar{x} = .2617$	$\bar{x} = .3076$	$\bar{x} = .3723$	$\bar{x}_M = .3111$
	$n = 7$	$n = 6$	$n = 6$	$n = 19$
Females	$\bar{x} = .2908$	$\bar{x} = .3891$	$\bar{x} = .3940$	$\bar{x}_F = .3631$
	$n = 6$	$n = 7$	$n = 8$	$n = 21$
	$\bar{x}_{RC} = .2751$	$\bar{x}_{LC} = .3514$	$\bar{x}_{SFS} = .3848$	
	$n = 13$	$n = 13$	$n = 14$	

Figure 1  
 Procedure X Format X Sex Interaction  
 For The GADU Variable





# References

Coombs, C. H. Portfolio theory and the measurement of risk. In M. F. Kaplan & S. Schwartz, Human judgment and decision processes. New York: Academic Press, Inc., 1975, 63-85.

Fischhoff, B., Slovic, P., & Lichtenstein, S. Fault trees: Sensitivity of estimated failure probabilities to problem representation. Technical report PTR-1042-77-8, 1977.

Fischhoff, B., Slovic, P., & Lichtenstein, S. Knowing what you want: Measuring labile values. Appearing in T. Wallsten (Ed.), Cognitive Processes in Choice and Decision Behavior. Hillsdale, New Jersey: Erlbaum, in press.

Hogarth, R.M. Cognitive processes and the assessment of subjective probability distributions. Journal of the American Statistical Association. 1975, 70, 350, 271-289.

Issacs, G. L., & Novick, M. R. Computer assisted data analysis monitor, 1978.

Kahneman, D., & Tversky, A. Prospect theory: An analysis of decision under risk. Econometrica. 1979, 47, 263-291.

Kahneman, D., & Tversky, A. Subjective probability: A judgment of representativeness. Cognitive Psychology. 1972, 3, 430-454.

Keeney, R. L., & Raiffa, H. Decisions with multiple objectives: Preferences and value tradeoffs. New York: John Wiley & Sons, 1976.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. Foundations of Measurement, Vol. I. New York: Academic Press, 1971.

Mosteller, F., & Nogee, P. An experimental measurement of utility. Journal of Political Economy. 1951, 59, 5, 371-404.

McDaniel, E. D., Halter, A. N., & Hartford, D. L. Grade utility: a new noncognitive factor in academic prediction. Educational and Psychological Measurement. 1961, 21, 3, 621-627.

- Novick, M. R., Chuang, D., & DeKeyser, D. Local and regional coherence utility assessment procedures. Trabajos de Estadística, in press, 1981.
- Novick, M. R., Isaacs, G. L., Hamer, R., Chen, J., Chuang, D., Woodworth, G., Molenaar, I., Lewis, C., & Libby, D. Computer-assisted data analysis monitor. 1980.
- Novick, M. R., & Lindley, D. V. Fixed-state assessment of utility functions. The Journal of the American Statistical Association, 1979, 74, 306-311.
- Novick, M. R., & Lindley, D. V. The use of more realistic utility functions in educational applications. Journal of Educational Measurement. 1978, 15, 181-192.
- Payne, J. W. Alternative approaches to decision making under risk: moments versus risk dimensions. Psychological Bulletin. 1973, 80, 6, 439-453.
- Raiffa, H., & Schlaifer, R. Applied statistical decision theory. Boston: Graduate School of Business Administration Harvard University, 1961.
- Schlaifer, R. Computer programs for elementary decision analysis. Boston: Graduate School of Business Administration Harvard University, 1971.
- Schlaifer, R. Probability and statistics for business decisions: An introduction to managerial economics under uncertainty. New York: McGraw-Hill Book Co., Inc., 1959.
- Slovic, P. From Shakespeare to Simon: Speculation - and some evidence - about man's ability to process information. Oregon Research Institute Research Bulletin. 1972, 12, (2) 1-29.
- Slovic, P., Fischhoff, B., & Lichtenstein, S. Behavioral decision theory. Annual Review of Psychology. 1977, 28, 1-39.
- Spetzler, C. S., & Stahl von Holstein, C-A. S. Probability encoding in decision analysis. Management Science. 1975, 22, 340-358.
- Swalm, R. O. Utility theory--insights into risk taking. Harvard Business Review. 1966, 44, 6, 123-136.

- Torgerson, W. S. Theory and methods of scaling. New York: John Wiley & Sons, Inc., 1958.
- Tversky, A. On the elicitation of preferences: Descriptive and prescriptive considerations. In D. E. Bell, R. L. Keeney, & H. Raiffa (Eds.), Conflicting objectives in decisions. Chichester: John Wiley & Sons, 1977, 209-222.
- Tversky, A., & Kahneman, D. Judgment under uncertainty: Heuristics and biases. Science. 1974, 185, 1124-1131.
- Vertinsky, I., & Wong, E. Eliciting preferences and the construction of indifference maps: a comparative empirical evaluation of two measurement methodologies. Socio-Economic Planning Sciences. 1975, 9, 15-24.
- Vlek, C.A.J. Coherence of Human Judgment in a limited probabilistic environment. Organizational Behavior and Human Performance. 1973, 9, 460-481.
- von Neumann, J., & Morgenstern, O. Theory of games and economic behavior. Princeton: Princeton University Press, 1953.

## Navy

- 1 Dr. Ed Aiken  
Navy Personnel R&D Center  
San Diego, CA 92152
- 1 Dr. Jack R. Borsting  
Provost & Academic Dean  
U.S. Naval Postgraduate School  
Monterey, CA 93940
- 1 Dr. Robert Breaux  
Code N-711  
NAVTRAEQUIPCEN  
Orlando, FL 32813
- 1 Chief of Naval Education and Training  
Liason Office  
Air Force Human Resource Laboratory  
Flying Training Division  
WILLIAMS AFB, AZ 85224
- 1 Deputy Assistant Secretary of the Navy  
(Manpower)  
Office of the Assistant Secretary of  
the Navy (Manpower, Reserve Affairs, 1  
and Logistics)  
Washington, DC 20350
- 1 Dr. Richard Elster  
Department of Administrative Sciences  
Naval Postgraduate School  
Monterey, CA 93940
- 1 DR. PAT FEDERICO  
NAVY PERSONNEL R&D CENTER  
SAN DIEGO, CA 92152
- 1 Mr. Paul Foley  
Navy Personnel R&D Center  
San Diego, CA 92152
- 1 Dr. John Ford  
Navy Personnel R&D Center  
San Diego, CA 92152
- 1 Dr. Henry M. Halff  
Department of Psychology, C-009  
University of California at San Diego  
La Jolla, CA 92093

## Navy

- 1 Dr. Patrick R. Harrison  
Psychology Course Director  
LEADERSHIP & LAW DEPT. (7b)  
DIV. OF PROFESSIONAL DEVELOPMENT  
U.S. NAVAL ACADEMY  
ANNAPOLIS, MD 21402
- 1 Dr. Norman J. Kerr  
Chief of Naval Technical Training  
Naval Air Station Memphis (75)  
Millington, TN 38054
- 1 Dr. William L. Maloy  
Principal Civilian Advisor for  
Education and Training  
Naval Training Command, Code 00A  
Pensacola, FL 32508
- 1 Dr. Kneale Marshall  
Scientific Advisor to DCNO(NPT)  
OP01T  
Washington DC 20370
- CAPT Richard L. Martin, USN  
Prospective Commanding Officer  
USS Carl Vinson (CVN-70)  
Newport News Shipbuilding and Drydock Co  
Newport News, VA 23607
- 1 Dr. James McBride  
Navy Personnel R&D Center  
San Diego, CA 92152
- 1 Dr William Montague  
Navy Personnel R&D Center  
San Diego, CA 92152
- 1 Ted M. I. Yellen  
Technical Information Office, Code 201  
NAVY PERSONNEL R&D CENTER  
SAN DIEGO, CA 92152
- 1 Library, Code P201L  
Navy Personnel R&D Center  
San Diego, CA 92152

## Navy

- 6 Commanding Officer  
Naval Research Laboratory  
Code 2627  
Washington, DC 20390
- 1 Psychologist  
ONR Branch Office  
Bldg 114, Section D  
666 Summer Street  
Boston, MA 02210
- 1 Psychologist  
ONR Branch Office  
536 S. Clark Street  
Chicago, IL 60605
- 1 Office of Naval Research  
Code 437  
800 N. Quincy Street  
Arlington, VA 22217
- 5 Personnel & Training Research Programs  
(Code 458)  
Office of Naval Research  
Arlington, VA 22217
- 1 Psychologist  
ONR Branch Office  
1030 East Green Street  
Pasadena, CA 91101
- 1 Office of the Chief of Naval Operations  
Research Development & Studies Branch  
(OP-115)  
Washington, DC 20350
- 1 LT Frank C. Petho, MSC, USN (Ph.D)  
Code L51  
Naval Aerospace Medical Research Laborat  
Pensacola, FL 32508
- 1 Mr. Arnold Rubenstein  
Naval Personnel Support Technology  
Naval Material Command (08T244)  
Room 1044, Crystal Plaza #5  
2221 Jefferson Davis Highway  
Arlington, VA 20360

## Navy

- 1 Dr. Robert G. Smith  
Office of Chief of Naval Operations  
OP-987H  
Washington, DC 20350
- 1 Dr. Alfred F. Smode  
Training Analysis & Evaluation Group  
(TAEG)  
Dept. of the Navy  
Orlando, FL 32813
- 1 Dr. Richard Sorensen  
Navy Personnel R&D Center  
San Diego, CA 92152
- 1 W. Gary Thomson  
Naval Ocean Systems Center  
Code 7132  
San Diego, CA 92152
- 1 Dr. Ronald Weitzman  
Code 54 WZ  
Department of Administrative Sciences  
U. S. Naval Postgraduate School  
Monterey, CA 93940
- 1 DR. MARTIN F. WISKOFF  
NAVY PERSONNEL R & D CENTER  
SAN DIEGO, CA 92152

## Army

- 1 Technical Director  
U. S. Army Research Institute for the  
Behavioral and Social Sciences  
5001 Eisenhower Avenue  
Alexandria, VA 22333
- 1 HQ USAREUE & 7th Army  
ODCSOPS  
USAREUE Director of GED  
APO New York 09403
- 1 DR. RALPH DUSEK  
U.S. ARMY RESEARCH INSTITUTE  
5001 EISENHOWER AVENUE  
ALEXANDRIA, VA 22333
- 1 Dr. Myron Fischl  
U.S. Army Research Institute for the  
Social and Behavioral Sciences  
5001 Eisenhower Avenue  
Alexandria, VA 22333
- 1 Dr. Harold F. O'Neil, Jr.  
Attn: PERI-OK  
Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333
- 1 Mr. Robert Ross  
U.S. Army Research Institute for the  
Social and Behavioral Sciences  
5001 Eisenhower Avenue  
Alexandria, VA 22333
- 1 Dr. Robert Sasmor  
U. S. Army Research Institute for the  
Behavioral and Social Sciences  
5001 Eisenhower Avenue  
Alexandria, VA 22333
- 1 Commandant  
US Army Institute of Administration  
Attn: Dr. Sherrill  
FT Benjamin Harrison, IN 46256
- 1 Dr. Frederick Steinheiser  
U. S. Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

## Army

- 1 Dr. Joseph Ward  
U.S. Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

## Air Force

- 1 Air Force Human Resources Lab  
AFHRL/MPD  
Brooks AFB, TX 78235
- 1 Air University Library  
AUL/LSE 76/443  
Maxwell AFB, AL 36112
- 1 Dr. Earl A. Alluisi  
HQ, AFHRL (AFSC)  
Brooks AFB, TX 78235
- 1 Dr. Genevieve Haddad  
Program Manager  
Life Sciences Directorate  
AFOSR  
Bolling AFB, DC 20332
- 1 Dr. Ross L. Morgan (AFHRL/LR)  
Wright -Patterson AFB  
Ohio 45433
- 1 Research and Measurment Division  
Research Branch, AFMPC/MPCYPR  
Randolph AFB, TX 78148
- 1 Dr. Malcolm Ree  
AFHRL/MP  
Brooks AFB, TX 78235
- 1 Dr. Marty Rockway  
Technical Director  
AFHRL(OT)  
Williams AFB, AZ 58224
- 1 Jack A. Thorp, Maj., USAF  
Life Sciences Directorate  
AFOSR  
Bolling AFB, DC 20332

## Marines

- 1 H. William Greenup  
Education Advisor (E031)  
Education Center, MCDEC  
Quantico, VA 22134
- 1 Director, Office of Manpower Utilization  
HQ, Marine Corps (MPU)  
BCB, Bldg. 2009  
Quantico, VA 22134
- 1 DR. A.L. SLAFKOSKY  
SCIENTIFIC ADVISOR (CODE RD-1)  
HQ, U.S. MARINE CORPS  
WASHINGTON, DC 20380

## CoastGuard

- 1 Mr. Thomas A. Warm  
U. S. Coast Guard Institute  
P. O. Substation 18  
Oklahoma City, OK 73169

## Other DoD

- 12 Defense Technical Information Center  
Cameron Station, Bldg 5  
Alexandria, VA 22314  
Attn: TC
- 1 Dr. Dexter Fletcher  
ADVANCED RESEARCH PROJECTS AGENCY  
1400 WILSON BLVD.  
ARLINGTON, VA 22209
- 1 Dr. William Graham  
Testing Directorate  
MEPCOM/MEPCT-P  
Ft. Sheridan, IL 60037
- 1 Military Assistant for Training and  
Personnel Technology  
Office of the Under Secretary of Defense  
for Research & Engineering  
Room 3D129, The Pentagon  
Washington, DC 20301
- 1 MAJOR Wayne Sellman, USAF  
Office of the Assistant Secretary  
of Defense (MRA&L)  
3B930 The Pentagon  
Washington, DC 20301
- 1 HEAD, SECTION ON MEDICAL EDUCATION  
UNIFORMED SERVICES UNIV. OF THE  
HEALTH SCIENCES  
6917 ARLINGTON ROAD  
BETHESDA, MD 20014



## Civil Govt

- 1 Dr. Susan Chipman  
Learning and Development  
National Institute of Education  
1200 19th Street NW  
Washington, DC 20208
- 1 Dr. Joseph I. Lipson  
SEDR W-638  
National Science Foundation  
Washington, DC 20550
- 1 Dr. John Mays  
National Institute of Education  
1200 19th Street NW  
Washington, DC 20208
- 1 Dr. Arthur Melmed  
National Institute of Education  
1200 19th Street NW  
Washington, DC 20208
- 1 Dr. Andrew R. Molnar  
Science Education Dev.  
and Research  
National Science Foundation  
Washington, DC 20550
- 1 Personnel R&D Center  
Office of Personnel Management  
1900 E Street NW  
Washington, DC 20415
- 1 Dr. Vern W. Urry  
Personnel R&D Center  
Office of Personnel Management  
1900 E Street NW  
Washington, DC 20415
- 1 Dr. Frank Withrow  
U. S. Office of Education  
400 Maryland Ave. SW  
Washington, DC 20202
- 1 Dr. Joseph L. Young, Director  
Memory & Cognitive Processes  
National Science Foundation  
Washington, DC 20550

## Non Govt

- 1 Dr. Erling B. Andersen  
Department of Statistics  
Studiestraede 6  
1455 Copenhagen  
DENMARK
- 1 1 psychological research unit  
Dept. of Defense (Army Office)  
Campbell Park Offices  
Canberra ACT 2600, Australia
- 1 Dr. Isaac Bejar  
Educational Testing Service  
Princeton, NJ 08450
- 1 Dr. Werner Birke  
DezWPs im Streitkraefteamt  
Postfach 20 50 03  
D-5300 Bonn 2  
WEST GERMANY
- 1 Dr. Nicholas A. Bond  
Dept. of Psychology  
Sacramento State College  
600 Jay Street  
Sacramento, CA 95819
- 1 Dr. David G. Bowers  
Institute for Social Research  
University of Michigan  
P. O. Box 1248  
Ann Arbor, MI 48106
- 1 Dr. Robert Brennan  
American College Testing Programs  
P. O. Box 168  
Iowa City, IA 52240
- 1 DR. C. VICTOR BUNDERSON  
WICAT INC.  
UNIVERSITY PLAZA, SUITE 10  
1160 SO. STATE ST.  
OREM, UT 84057
- 1 Dr. John B. Carroll  
Psychometric Lab  
Univ. of No. Carolina  
Davie Hall 013A  
Chapel Hill, NC 27514

## Non Govt

- 1 Charles Myers Library  
Livingstone House  
Livingstone Road  
Stratford  
London E15 2LJ  
ENGLAND
- 1 Dr. Kenneth E. Clark  
College of Arts & Sciences  
University of Rochester  
River Campus Station  
Rochester, NY 14627
- 1 Dr. Norman Cliff  
Dept. of Psychology  
Univ. of So. California  
University Park  
Los Angeles, CA 90007
- 1 Dr. William E. Coffman  
Director, Iowa Testing Programs  
334 Lindquist Center  
University of Iowa  
Iowa City, IA 52242
- 1 Dr. Meredith P. Crawford  
American Psychological Association  
1200 17th Street, N.W.  
Washington, DC 20036
- 1 Dr. Hans Crombag  
Education Research Center  
University of Leyden  
Boerhaavelsaan 2  
2334 EN Leyden  
The NETHERLANDS
- 1 Dr. Emmanuel Donchin  
Department of Psychology  
University of Illinois  
Champaign, IL 61820
- 1 LCOL J. C. Eggenberger  
DIRECTORATE OF PERSONNEL APPLIED RESEARCH  
NATIONAL DEFENCE HQ  
101 COLONEL BY DRIVE  
OTTAWA, CANADA K1A 0K2

## Non Govt

- 1 Dr. Leonard Feldt  
Lindquist Center for Measurement  
University of Iowa  
Iowa City, IA 52242
- 1 Dr. Richard L. Ferguson  
The American College Testing Program  
P.O. Box 168  
Iowa City, IA 52240
- 1 Dr. Victor Fields  
Dept. of Psychology  
Montgomery College  
Rockville, MD 20850
- 1 Univ. Prof. Dr. Gerhard Fischer  
Liebiggasse 5/3  
A 1010 Vienna  
AUSTRIA
- 1 Professor Donald Fitzgerald  
University of New England  
Armidale, New South Wales 2351  
AUSTRALIA
- 1 Dr. Edwin A. Fleishman  
Advanced Research Resources Organ.  
Suite 900  
4330 East West Highway  
Washington, DC 20014
- 1 Dr. John R. Frederiksen  
Bolt Beranek & Newman  
50 Moulton Street  
Cambridge, MA 02138
- 1 DR. ROBERT GLASER  
LRDC  
UNIVERSITY OF PITTSBURGH  
3939 O'HARA STREET  
PITTSBURGH, PA 15213
- 1 Dr. Ron Hambleton  
School of Education  
University of Massachusetts  
Amherst, MA 01002

## Non Govt

- 1 Dr. Chester Harris  
School of Education  
University of California  
Santa Barbara, CA 93106
- 1 Dr. Lloyd Humphreys  
Department of Psychology  
University of Illinois  
Champaign, IL 61820
- 1 Library  
HUMRRO/Western Division  
27857 Berwick Drive  
Carmel, CA 93921
- 1 Dr. Steven Hunka  
Department of Education  
University of Alberta  
Edmonton, Alberta  
CANADA
- 1 Dr. Earl Hunt  
Dept. of Psychology  
University of Washington  
Seattle, WA 98105
- 1 Dr. Huynh Huynh  
College of Education  
University of South Carolina  
Columbia, SC 29208
- 1 Dr. Douglas H. Jones  
Rm T-255  
Educational Testing Service  
Princeton, NJ 08450
- 1 Professor John A. Keats  
University of Newcastle  
AUSTRALIA 2308
- 1 Dr. Mazie Knerr  
Litton-Mellonics  
Box 1286  
Springfield, VA 22151
- 1 Mr. Marlin Kroger  
1117 Via Goleta  
Palos Verdes Estates, CA 90274

## Non Govt

- 1 Dr. Michael Levine  
Department of Educational Psychology  
210 Education Bldg.  
University of Illinois  
Champaign, IL 61801
- 1 Dr. Charles Lewis  
Faculteit Sociale Wetenschappen  
Rijksuniversiteit Groningen  
Oude Boteringestraat  
Groningen  
NETHERLANDS
- 1 Dr. Robert Linn  
College of Education  
University of Illinois  
Urbana, IL 61801
- 1 Dr. Frederick M. Lord  
Educational Testing Service  
Princeton, NJ 08540
- 1 Dr. Gary Marco  
Educational Testing Service  
Princeton, NJ 08450
- 1 Dr. Scott Maxwell  
Department of Psychology  
University of Houston  
Houston, TX 77004
- 1 Dr. Samuel T. Mayo  
Loyola University of Chicago  
820 North Michigan Avenue  
Chicago, IL 60611
- 1 Dr. James A. Paulson  
Portland State University  
P.O. Box 751  
Portland, OR 97207
- 1 MR. LUIGI PETRULLO  
2431 N. EDGEWOOD STREET  
ARLINGTON, VA 22207
- 1 DR. DIANE M. RAMSEY-KLEE  
R-K RESEARCH & SYSTEM DESIGN  
3947 RIDGEMONT DRIVE  
MALIBU, CA 90265

## Non Govt

- 1 MINRAT H. L. RAUCH  
P II 4  
BUNDESMINISTERIUM DER VERTEIDIGUNG  
POSTFACH 1328  
D-53 BONN 1, GERMANY
- 1 Dr. Mark D. Reckase  
Educational Psychology Dept.  
University of Missouri-Columbia  
4 Hill Hall  
Columbia, MO 65211
- 1 Dr. Andrew M. Rose  
American Institutes for Research  
1055 Thomas Jefferson St. NW  
Washington, DC 20007
- 1 Dr. Leonard L. Rosenbaum, Chairman  
Department of Psychology  
Montgomery College  
Rockville, MD 20850
- 1 Dr. Ernst Z. Rothkopf  
Bell Laboratories  
600 Mountain Avenue  
Murray Hill, NJ 07974
- 1 Dr. Lawrence Rudner  
403 Elm Avenue  
Takoma Park, MD 20012
- 1 Dr. J. Ryan  
Department of Education  
University of South Carolina  
Columbia, SC 29208
- 1 PROF. FUMIKO SAMEJIMA  
DEPT. OF PSYCHOLOGY  
UNIVERSITY OF TENNESSEE  
KNOXVILLE, TN 37916
- 1 Committee on Cognitive Research  
1 Dr. Lonnie R. Sherrod  
Social Science Research Council  
605 Third Avenue  
New York, NY 10016

## Non Govt

- 1 Dr. Kazuo Shigematsu  
University of Tohoku  
Department of Educational Psychology  
Kawauchi, Sendai 980  
JAPAN
- 1 Dr. Richard Snow  
School of Education  
Stanford University  
Stanford, CA 94305
- 1 Dr. Robert Sternberg  
Dept. of Psychology  
Yale University  
Box 11A, Yale Station  
New Haven, CT 06520
- 1 DR. PATRICK SUPPES  
INSTITUTE FOR MATHEMATICAL STUDIES IN  
THE SOCIAL SCIENCES  
STANFORD UNIVERSITY  
STANFORD, CA 94305
- 1 Dr. Hariharan Swaminathan  
Laboratory of Psychometric and  
Evaluation Research  
School of Education  
University of Massachusetts  
Amherst, MA 01003
- 1 Dr. Brad Symson  
Psychometric Research Group  
Educational Testing Service  
Princeton, NJ 08541
- 1 Dr. Kikumi Tatsuoka  
Computer Based Education Research  
Laboratory  
252 Engineering Research Laboratory  
University of Illinois  
Urbana, IL 61801
- 1 Dr. David Thissen  
Department of Psychology  
University of Kansas  
Lawrence, KS 66044

## Non Govt

- 1 Dr. J. Uhlaner  
Perceptronics, Inc.  
6271 Variel Avenue  
Woodland Hills, CA 91364
- 1 Dr. Howard Wainer  
Bureau of Social Science Research  
1990 M Street, N. W.  
Washington, DC 20036
- 1 DR. THOMAS WALLSTEN  
PSYCHOMETRIC LABORATORY  
DAVIE HALL 013A  
UNIVERSITY OF NORTH CAROL  
CHAPEL HILL, NC 27514
- 1 Dr. David J. Weiss  
N660 Elliott Hall  
University of Minnesota  
75 E. River Road  
Minneapolis, MN 55455
- 1 DR. SUSAN E. WHITELY  
PSYCHOLOGY DEPARTMENT  
UNIVERSITY OF KANSAS  
LAWRENCE, KANSAS 66044
- 1 Wolfgang Wildgrube  
Streitkræfteamt  
Box 20 50 03  
D-5300 Bonn 2  
WEST GERMANY
- 1 Dr. J. Arthur Woodward  
Department of Psychology  
University of California  
Los Angeles, CA 90024
- 1 Dr. Karl Zinn  
Center for research on Learning  
and Teaching  
University of Michigan  
Ann Arbor, MI 48104